Radiotherapy treatment planning for head and neck cancer: Automation and improved organ at risk sparing

Tol, J.P.

2016

document version
Publisher's PDF, also known as Version of record

Radiotherapy treatment planning for head and neck cancer: Automation and improved organ at risk sparing

Jim Pieter Tol

Source cover image: Bonamy, Paris 1860, MLI VUmc

Printed by: Ipskamp Printing B.V., Enschede

ISBN: 978-94-028-0257-3

VRIJE UNIVERSITEIT

Radiotherapy treatment planning for head and neck cancer: Automation and improved organ at risk sparing

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op vrijdag 9 september 2016 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Jim Pieter Tol
geboren te Den Helder

promotor: prof.dr. B.J. Slotman

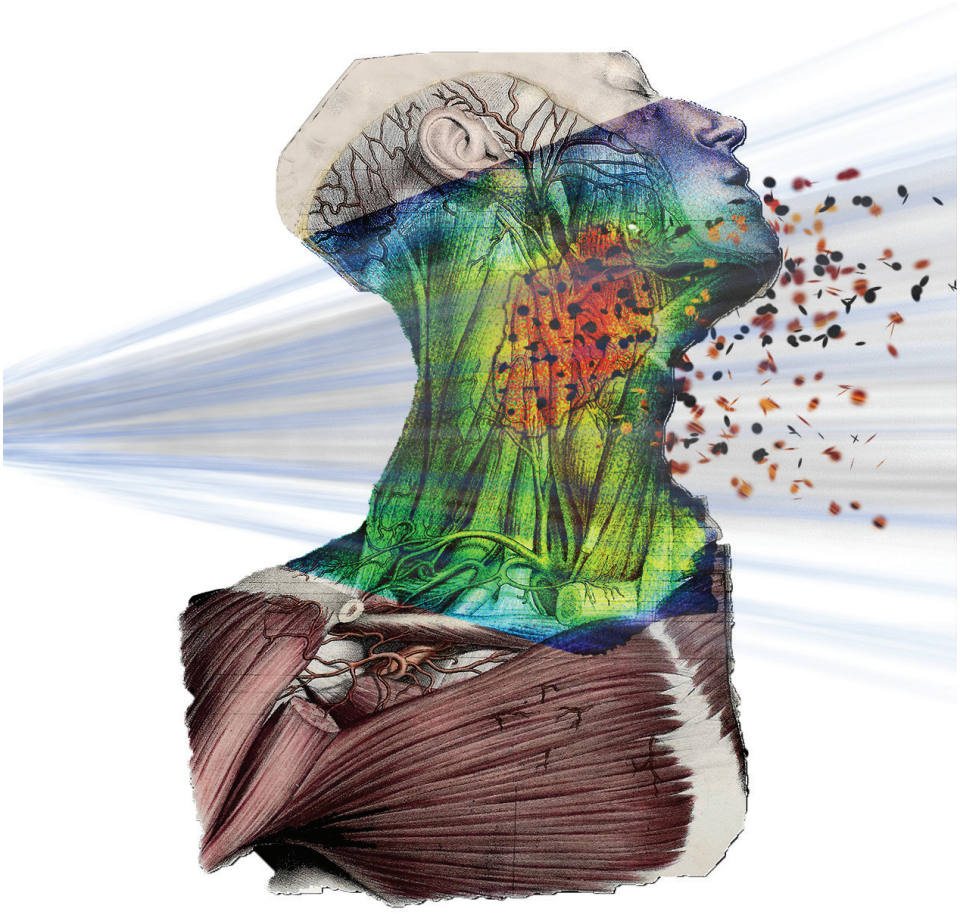
copromotoren: dr.ir. W.F.A.R. Verbakel

dr. M.R. Dahele

Contents

Chapter 1

Introduction



Introduction

Cancer has been identified by the World Health Organization as the second leading cause of death worldwide[1], exceeded only by cardiovascular diseases, and resulted in over 8 million deaths worldwide in 2012[2]. This amounts to roughly 15% of all deaths globally. The burden is even higher in the developed world[3], with cancer accounting for approximately 25% of all disease related deaths[1]. Over 14 million new cancer cases were identified worldwide in 2012, and due to the general growth and increasing age of the population, this figure is expected to rise to 21.7 million in 2030, with 13 million cancer associated deaths[3]. The number of newly diagnosed cancer cases in the Netherlands has increased from approximately 57.000 in 1990, to over 100.000 in 2013, although this near doubling was associated with a proportionally lower increase of 7.500 deaths[4]. Cancer presents a substantial economic burden on society, accounting for 5.3% of the total health care costs in the Netherlands in 2011[5].

Three major treatment options for cancer are generally recognized as being surgery, radiotherapy and chemotherapy. The specific treatment chosen, or combination thereof, depends on many factors, including the site and stage of the disease and the patient's biological age, overall health and personal preference. Roughly 50% of patients will receive radiotherapy as part of their treatment[6], and it has been estimated to make an important contribution to the outcome in 40% of patients who are ultimately cured of their disease[7,8]. Despite this, radiotherapy treatments have in the past accounted for only 5% of the total cancer care costs[9], and are therefore considered to be highly cost-effective.

This thesis primarily revolves around improving and automating the radiotherapy treatment planning process for locally advanced head and neck cancer (HNC) patients. Although HNC is not the most common form of cancer, with roughly 3.000 cases identified yearly in the Netherlands[4], radiotherapy, frequently combined with chemotherapy, plays a key role in its treatment. In the majority of patients this combination is at least as effective as surgical resection and potentially curative, with the possibility of organ-preservation as an added advantage[10–12]. Organ-preservation is important to maximize the functional outcome of patients after treatment (such as retaining the ability to swallow) and allowing them to maintain their cosmetic appearance. As a result, concurrent chemo-radiotherapy (CRT) has become the standard of care for patients with locally advanced HNC[13]. However, side-effects of CRT for HNC, including chronic dry mouth (xerostomia), swallowing dysfunction (dysphagia) and painful inflammation and ulceration of the mucous membranes that line the digestive tract (oral mucositis) can place a burden on the quality-of-life after treatment.

The probability of developing organ dysfunction has been correlated with radiation dose for a number of these toxicities[14–17].

Radiotherapy involves the use of ionizing radiation to control or kill malignant cells. The first step in the radiotherapy process for HNC is acquiring a computed tomography (CT) scan in treatment position and delineation of the gross tumor volume (GTV) that is represented by the visible tumor on clinical examination and on co-registered scans, including magnetic resonance imaging (MRI) and positron emission tomography (PET) / CT. The clinical target volume (CTV) is created by the inclusion of a margin around the GTV to account for microscopic spread of the disease. To account for additional geometric inaccuracies, such as uncertainty in patient positioning, a final margin is added around the CTV resulting in the planning target volume (PTV). In HNC treatments, the CTV and PTV are often divided into boost and elective regions. The former are treated to a higher dose because they typically include biopsy-proven positive or imaging positive tumor regions, while the latter solely include elective lymph node regions, i.e., areas in the head and neck region to which the tumor cells are likely to spread.

Since the ionizing radiation must travel through the body to reach the target volume, healthy tissues are inevitably also irradiated and may be damaged by radiotherapy treatments. The use of multiple gantry angles can disperse the radiation doses over the body and may reduce the amount of dose being delivered to radiosensitive organs-at-risk (OARs), while the summation of the individual treatment fields in the PTV ensures that sufficiently high doses are given to the tumor region. Shaping individual treatment fields using a multileaf collimator (MLC) may further decrease OAR doses. An MLC consists of a large number of metal leaves (typically made from tungsten) that can move individually. By opening these leaves such that the aperture resembles the shape of the tumor as seen from the beam's eye view of each treatment field, the tissues surrounding the PTVs are shielded from radiation, reducing OAR doses. This treatment technique forms the basis of three-dimensional conformal radiotherapy (3D-CRT)[18].

During 3D-CRT, uniform radiation doses are typically delivered by each field using fixed MLC leaf configurations and flattened radiation beams, although the relative contribution to the total delivered dose can vary per field. Intensity modulated radiotherapy (IMRT)[12,13] is a more advanced form of radiation therapy delivery and uses programmed movement of the MLC leaves to deliver inhomogeneous dose distributions per treatment field. Only the summation of each individual field results in relatively homogeneous and sufficiently high doses to the tumor region. IMRT can often substantially improve OAR sparing over

3D-CRT in the high dose range, particularly for OARs that are located in close proximity to the tumor, or for OARs that are surrounded by concave tumors. Volumetric modulated arc therapy (VMAT) is a more recent evolution of IMRT in which continuous movement of the gantry is combined with simultaneous movements of the individual MLC leaves[21] and varying dose rates, often resulting in a shorter total treatment time compared to static-field IMRT. In addition, VMAT has been shown to reduce monitor unit (MU) requirements for some disease sites[22]. Studies benchmarking VMAT and IMRT plans for HNC have shown similar plan quality in terms of PTV dose homogeneity and OAR sparing[23,24].

3D-CRT treatment plans are created through a so-called forward treatment planning process in which the planner manually determines appropriate gantry angles, collimator angles, field sizes, MLC leaf configurations and relative contributions of each field to the total delivered dose for individual patients. For IMRT and VMAT, optimal MLC settings cannot be determined in a straightforward fashion and instead rely on the use of optimization algorithms to determine the leaf configurations. During the optimization phase the planner typically places objectives on delineated structures to which the resulting plan should comply, including, for example, minimum dose requirements for the PTVs and maximum dose requirements for critical OARs. Through minimization of a cost function in which the relative contributions of all dose-volume objectives are weighted[25], the optimization algorithm attempts to determine the MLC leaf configuration that provides an optimal balance between satisfying the various objectives. However, optimally achievable levels of individual OAR sparing can differ widely between patients as their geometrical features (such as the location of the tumor with respect to the OARs) are not constant[26]. Using a fixed set of optimization objectives may therefore prevent optimal plan quality from being reached for individual patients because the optimization objectives are either satisfied too easily, preventing further OAR dose reductions from being obtained, or because certain objectives cannot be satisfied without compromising others (for example leading to higher OAR doses or a deterioration in PTV dose homogeneity). Treatment planning for IMRT and VMAT therefore requires an iterative approach in which the planner evaluates the initial plan, changes the dose-volume objectives in an attempt to further modify the dose distributions, and restarts the optimization process. Each iteration ideally navigates closer to the best possible set of optimization objectives for each individual patient. Since the planner has to manually select appropriate objectives at each optimization run, this process is time-consuming and a major source of variability in the planning process, and has been shown to lead to suboptimal and inconsistent plan quality both between different radiotherapy clinics and planners within the same clinic[27–30]. This variability is exacerbated by the field

setup that has to be determined manually before the optimization process is started. In IMRT, beam angle optimization has been investigated[31–33] which could potentially remove this source of variability.

IMRT and VMAT have facilitated the shaping of increasingly complex dose distributions, allowing for the inclusion of more OARs for sparing. HNC planning evolved from solely attempting spinal cord and brainstem sparing, to include sparing of the parotid and submandibular glands, oral cavity and individual swallowing muscles, respectively aiming to reduce radiation induced levels of xerostomia, oral mucositis and dysphagia. This large number of OARs makes it difficult to objectively evaluate resulting plan quality as the trade-offs between achieving different planning goals are not always apparent. Furthermore, planning requirements may vary between radiotherapy institutes and treating clinicians, influencing the dose distributions of plans deemed acceptable. A plan is considered Pareto-optimal if no aspect of it can be improved without deteriorating another[34]. Although Pareto-optimal planning has been widely investigated in recent years, it is important to note that Pareto-optimality of a plan does not imply clinical optimality.

Automated planning solutions have been suggested to improve the consistency and quality of radiotherapy treatment plans. Two approaches to automate the optimization process have received the most attention. The first, often regarded as knowledge-based planning (KBP)[35–38], attempts to estimate the optimal placement of the initial optimization objectives, thereby removing the need for performing multiple subsequent optimization runs. KBP typically involves the creation of a model that consists of the geometric and dosimetric features of a large number of previously created treatment plans. This model can be used to predict achievable OAR doses for future patients, taking into account their unique geometrical characteristics. These predicted OAR doses can serve as patient-specific optimization objectives for the optimization of future plans. Based on in-house developed KBP solutions[26,35], Varian Medical Systems (Palo Alto, USA) released RapidPlan, a commercial approach to KBP, in 2014.

The second approach attempts to automatically navigate towards optimal placement of the optimization objectives. Multi-Criteria Optimization (MCO) has been the subject of numerous investigations in the previous years[39–42], and in-house developed solutions have been implemented in routine clinical practice with success. iCycle is one such solution[43–45] and uses a wish-list specifying various optimization goals to automatically steer the subsequently performed optimizations towards the desired high-quality plan. Recent investigations regarding MCO navigation, a topic closely related to automated planning,

present the physician with a collection of previously determined Pareto-optimal plans (the Pareto-front)[27,28]. By navigating over this front the physician can select a plan with clinically optimal trade-offs regarding OAR sparing and PTV dose homogeneity. Pareto-fronts have also been used to benchmark different treatment planning techniques[47–50]. Pinnacle Auto-Planning is the commercial solution by Philips Radiation Oncology Systems, and uses an iterative approach to automatically adapt the optimization objectives, constraints and dose shaping contours during the optimization process[51]. These solutions automate the repetitive and time-consuming tasks that the planners have to perform manually, potentially freeing up time that can be invested in other activities. Automating the decision making at certain parts of the planning process may also contribute to improved planning consistency and quality.

Certain limitations to these automated planning strategies remain. A KBP model library, for example, needs to consist of high quality plans from a varying range of OAR-PTV geometries, while judging the quality of these plans can be elusive. MCO solutions that navigate to a single resulting plan require optimization aims (or wish-list), while it is not guaranteed that this list will result in clinically optimal plans for individual patients. MCO techniques that allow Pareto-front navigation still require the user to evaluate a large number of planning trade-offs in detail before deciding on a final treatment plan. Finally, assessment of plan quality still relies on human judgment, meaning that the automated solutions need to be augmented by high-level human expertise in treatment planning and judging treatment plans.

Outline of this thesis

The links between radiation dosage and clinically relevant toxicities in multiple organs, make HNC an ideal paradigm for studying advanced radiotherapy treatment planning. Soon after introducing VMAT for HNC patients, it was discovered that a dual-arc technique, irradiating using a clockwise and counter-clockwise rotation of the gantry, can offer improvements in plan quality over single arc VMAT plans. In **Chapter 2** it was systematically investigated whether further gains in plan quality are feasible when using more than two arcs in the treatment planning process.

Chapter 3 explores the trade-offs between PTV dose homogeneity and OAR sparing that may occur during optimization when increasing the optimization weightings on the PTV optimization objectives. A straightforward KBP method is presented and used to predict patient-specific levels of OAR sparing at constant PTV dose homogeneity levels. The clinical implications of these trade-offs are further explored in **Chapter 4** by creating

plans that satisfy varying PTV dose coverage and homogeneity requirements specified by three different planning protocols. Varying levels of OAR sparing resulting from this comparison are particularly important when comparing the results of clinical trials, as they may influence normal tissue complication probabilities (NTCPs).

Our in-house developed automatic interactive optimizer (AIO) has been used in routine clinical practice since February 2014 to automate the optimization phase of HNC treatment planning. During the interactive optimization process of the Eclipse TPS, AIO gradually lowers the position of the OAR optimization objectives, thereby ensuring that adequate attempts to OAR sparing are made at all stages during optimization. **Chapter 5** provides an account of the technical development of AIO, and explains how the solution works in conjunction with the VMAT optimization algorithm. Using ten HNC patients, the influence of various AIO settings on resulting plan quality is investigated, and benchmarked against manually optimized plans. In **Chapter 6,** optimal AIO settings resulting from this preliminary investigation are used to provide a detailed clinical investigation of AIO plan quality on a large set of HNC patients, including a blinded comparison by a head and neck radiation oncologist. The use of AIO for twenty locally advanced lung cancer patients is also investigated to demonstrate the versatility of AIO for different treatment sites.

Chapter 7 investigated the performance of RapidPlan using a sixty patient HNC model. Resulting plans were benchmarked against recent HNC plans. An additional evaluation group consisted of patients that were treated shortly after RapidArc VMAT was introduced in our department, demonstrating the effect of increasing planner skill and experience on resulting plan quality. Dosimetric outliers in the form of sub-optimal plans may be present in KBP models. Such outliers can influence the correlation between the geometric and dosimetric features of the model, and may influence the resulting plan quality. Using varying numbers of dosimetric outliers, **Chapter 8** investigates the effect on resulting plan quality for HNC. In **Chapter 9**, the prediction accuracy of a ninety patient HNC model was investigated by determining the similarity between the predicted OAR dose-volume histograms (DVHs), and the DVHs that were obtained when these predictions were used to guide the optimization of a new treatment plan. Accurate DVH predictions may allow KBP models to be used as a rapid plan-quality evaluation tool.

Novel dose delivery and treatment planning techniques are expected to improve HNC plan quality. The expected improvements, however, are often based on traditional treatment planning studies that compare different planning techniques on the same set of patients. Once the preferred technique is introduced in clinical practice, the expected gains in

plan quality are rarely evaluated. **Chapter 10** therefore provides a detailed investigation of routine HNC plan quality resulting from the introduction of new dose delivery and treatment planning techniques in four subsequent periods during the past decade at the department of Radiation Oncology at the VU University Medical Center.

IMRT has in recent years established itself as the standard of care in HNC radiotherapy, while the use of VMAT is increasing[52]. In theory, proton therapy has the potential to further improve OAR sparing because the finite range of the protons causes a steep dose fall-off after their dose maximum (the Bragg-peak). **Chapter 11** benchmarks different proton therapy techniques against VMAT, simultaneously taking into account plan robustness (invariance of the proton dose distributions to range uncertainties and setup errors), OAR sparing and PTV dose homogeneity.

Chapter 12 concludes by discussing the research that is presented in this thesis in the context of current clinical practice, the published literature and future perspectives.

References

1 World Health Organization - International Agency for Research on Cancer, World Cancer Report 2014 (IARC, 2014).

2 American Cancer Society, Global Cancer Facts & Figures 3rd Edition. Atlanta: Am. Cancer Soc. (2015).

3 F. Bray, A. Jemal, N. Grey, J. Ferlay, and D. Forman, "Global cancer transitions according to the Human Development Index (2008-2030): a population-based study.," Lancet. Oncol. **13**(8), 790–801 (2012).

4 Integraal kankercentrum Nederland (IKNL), Dutch cancer registration, http://cijfersoverkanker.nl, accessed on Novermber 25th 2015.

5 Nationaal Kompas Volksgezondheid, 'Hoeveel zorg gebruiken mensen met kanker en wat zijn de kosten?' http://nationaalkompas.nl, accessed on November 25th 2015.

6 G. Delaney, S. Jacob, C. Featherstone, and M. Barton, "The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines.," Cancer **104**(6), 1129–37 (2005).

7 G.C. Barnett, C.M.L. West, A.M. Dunning, R.M. Elliott, C.E. Coles, P.D.P. Pharoah, and N.G. Burnet, "Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype.," Nat. Rev. Cancer **9**(2), 134–42 (2009).

8 R. Baskar, K.A. Lee, R. Yeo, and K.W. Yeoh, "Cancer and radiation therapy: current advances and future directions.," Int. J. Med. Sci. **9**(3), 193–9 (2012).

9 U. Ringborg, D. Bergqvist, B. Brorsson, E. Cavallin-ståhl, J. Ceberg, N. Einhorn, J. Frödin, J. Järhult, G. Lamnevik, C. Lindholm, B. Littbrand, A. Norlund, U. Nylén, M. Rosén, H. Svensson, and T.R. Möller, "The Swedish Council on Technology Assessment in Health Care (SBU) Systematic Overview of Radiotherapy for Cancer including a Prospective Survey of Radiotherapy Practice in Sweden 2001-Summary and Conclusions," Acta Oncol. (Madr). **42**(5-6), 357–365 (2003).

10 A.A. Forastiere, H. Goepfert, M. Maor, T.F. Pajak, R. Weber, W. Morrison, B. Glisson, A. Trotti, J.A. Ridge, C. Chao, G. Peters, D. Lee, A. Leaf, J. Ensley, and J. Cooper, "Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer.," N. Engl. J. Med. 349(22), 2091–8 (2003).

11 P. Dirix and S. Nuyts, "Evidence-based organ-sparing radiotherapy in head and neck cancer.," Lancet. Oncol. **11**(1), 85–91 (2010).

12 A.A. Forastiere, Q. Zhang, R.S. Weber, M.H. Maor, H. Goepfert, T.F. Pajak, W. Morrison, B. Glisson, A. Trotti, J. a Ridge, W. Thorstad, H. Wagner, J.F. Ensley, and J.S. Cooper, "Long-term results of RTOG 91-11: a comparison of three nonsurgical treatment strategies to preserve the larynx in patients with locally advanced larynx cancer.," J. Clin. Oncol. **31**(7), 845–52 (2013).

13 J.P. Pignon, A. le Maître, E. Maillard, and J. Bourhis, "Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients.," Radiother. Oncol. **92**(1), 4–14 (2009).

14 G. Sanguineti, M.P. Sormani, S. Marur, G.B. Gunn, N. Rao, M. Cianchetti, F. Ricchetti, T. McNutt, B. Wu, and A.A. Forastiere, "Effect of radiotherapy and chemotherapy on the risk of mucositis during intensity-modulated radiation therapy for oropharyngeal cancer.," Int. J. Radiat. Oncol. Biol. Phys. **83**(1), 235–42 (2012).

15 T. Dijkema, C.P.J. Raaijmakers, R.K. Ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

16 J.O. Deasy, V. Moiseenko, L. Marks, K.S.C. Chao, J. Nam, and A. Eisbruch, "Radiotherapy dose-volume effects on salivary gland function.," Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl), S58–63 (2010).

17 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F.A.M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

18 B.B. Mittal, J.A. Purdy, and K.K. Ang, Advances in Radiation Therapy (Springer US, Boston, MA, 1998).

19 A. Brahme, J.E. Roos, and I. Lax, "Solution of an integral equation encountered in rotation therapy.," Phys. Med. Biol. **27**(10), 1221–9 (1982).

20 T. Bortfeld, "IMRT: a review and preview.," Phys. Med. Biol. **51**(13), R363–79 (2006).

21 K. Otto, "Volumetric modulated arc therapy: IMRT in a single gantry arc.," Med. Phys. **35**(1), 310–7 (2008).

22 J.L. Bedford, "Treatment planning for volumetric modulated arc therapy," Med. Phys. **36**(11), 5128 (2009).

23 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–9 (2009).

24 M. Scorsetti, A. Fogliata, S. Castiglioni, C. Bressi, M. Bignardi, P. Navarria, P. Mancosu, L. Cozzi, S. Pentimalli, F. Alongi, and A. Santoro, "Early clinical experience with volumetric modulated arc therapy in head and neck cancer patients.," Radiat. Oncol. **5**(1), 93 (2010).

25 S. Webb, "Optimisation of conformal radiotherapy dose distributions by simulated annealing," **34**(10), 1349–1370 (1989).

26 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

27 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

28 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

29 S. Everitt, T. Kron, N. Fimmell, J. Reynolds, C. Laferlita, D. Ball, M. Schneider-Kolsky, R. Budd, and M. Macmanus, "Interplanner variability in carrying out three-dimensional conformal radiation therapy for non-small-cell lung cancer.," J. Med. Imaging Radiat. Oncol. **52**(3), 293–6 (2008).

30 Y. Matsuo, K. Takayama, Y. Nagata, E. Kunieda, K. Tateoka, N. Ishizuka, T. Mizowaki, Y. Norihisa, M. Sakamoto, Y. Narita, S. Ishikura, and M. Hiraoka, "Interinstitutional variations in planning for stereotactic body radiation therapy for lung cancer.," Int. J. Radiat. Oncol. Biol. Phys. **68**(2), 416–25 (2007).

31 S. Söderström and A. Brahme, "Which is the most suitable number of photon beam portals in coplanar radiation therapy?," Int. J. Radiat. Oncol. Biol. Phys. **33**(1), 151–9 (1995).

32 X. Wang, X. Zhang, L. Dong, H. Liu, Q. Wu, and R. Mohan, "Development of methods for beam angle optimization for IMRT using an accelerated exhaustive search strategy.," Int. J. Radiat. Oncol. Biol. Phys. **60**(4), 1325–37 (2004).

33 D. Craft, "Local beam angle optimization with linear programming and gradient search.," Phys. Med. Biol. **52**(7), N127–35 (2007).

34 V. Pareto, Cours d'Économie Politique (F. Pichou, Lausanne, 1897).

35 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

36 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

37 A. Fogliata, P.M. Wang, F. Belosi, A. Clivio, G. Nicolini, E. Vanetti, and L. Cozzi, "Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer.," Radiat. Oncol. **9**(1), 236 (2014).

38 A. Fogliata, F. Belosi, A. Clivio, P. Navarria, G. Nicolini, M. Scorsetti, E. Vanetti, and L. Cozzi, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer.," Radiother. Oncol. **113**(3), 385–91 (2014).

39 D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning.," Med. Phys. **33**(9), 3399–407 (2006).

40 D.L. Craft, T.S. Hong, H.A. Shih, and T.R. Bortfeld, "Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **82**(1), e83–90 (2012).

41 D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld, "Multicriteria VMAT optimization.," Med. Phys. **39**(2), 686–96 (2012).

42 T.S. Hong, D.L. Craft, F. Carlsson, and T.R. Bortfeld, "Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head.," Int. J. Radiat. Oncol. Biol. Phys. **72**(4), 1208–14 (2008).

43 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

44 P.W.J. Voet, S. Breedveld, M.L.P. Dirkx, P.C. Levendag, and B.J.M. Heijmen, "Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT.," Med. Phys. **39**(8), 4858–65 (2012).

45 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

46 D. Craft, P. Süss, and T. Bortfeld, "The tradeoff between treatment plan quality and required number of monitor units in intensity-modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **67**(5), 1596–605 (2007).

47 Z. van Kesteren, T.M. Janssen, E. Damen, and C. van Vliet-Vroegindeweij, "The dosimetric impact of leaf interdigitation and leaf width on VMAT treatment planning in Pinnacle: comparing Pareto fronts.," Phys. Med. Biol. **57**(10), 2943–52 (2012).

48 R.O. Ottosson, P.E. Engstrom, D. Sjöström, C.F. Behrens, A. Karlsson, T. Knöös, and C. Ceberg, "The feasibility of using Pareto fronts for comparison of treatment planning systems and delivery techniques.," Acta Oncol. **48**(2), 233–7 (2009).

49 R.O. Ottosson, A. Karlsson, and C.F. Behrens, "Pareto front analysis of 6 and 15 MV dynamic IMRT for lung cancer using pencil beam, AAA and Monte Carlo.," Phys. Med. Biol. **55**(16), 4521–33 (2010).

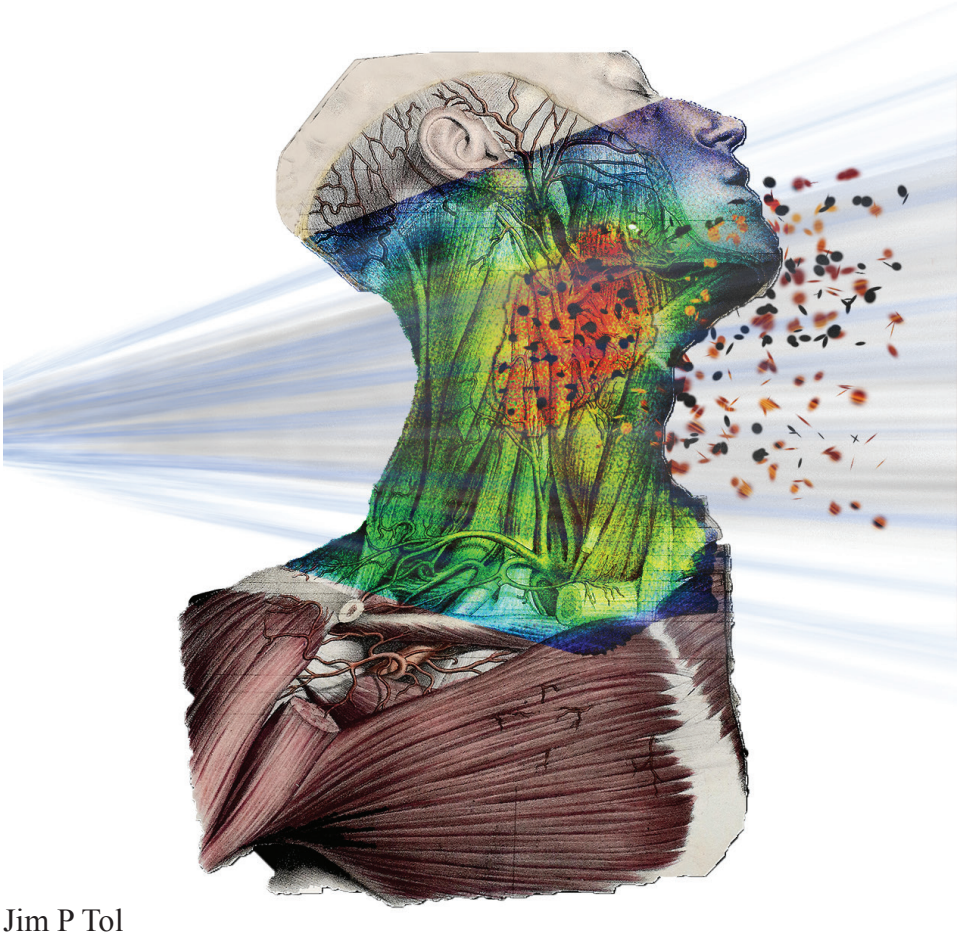
50 T. Janssen, Z. van Kesteren, G. Franssen, E. Damen, and C. van Vliet-Vroegindeweij, "Pareto fronts in clinical practice for pinnacle.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 873–80 (2013).

51 J. Krayenbuehl, I. Norton, G. Studer, and M. Guckenberger, "Evaluation of an automated knowledge based treatment planning system for head and neck.," Radiat. Oncol. **10**(1), 226 (2015).

52 American Society for Radiation Oncology, ASTRO Model Policies - Intensity Modulated Radiation Therapy (IMRT) (2015).

Chapter 2

Increasing the number of arcs improves head and neck volumetric modulated arc therapy plans



Jim P Tol
Max Dahele
Ben J Slotman
Wilko FAR Verbakel

Acta Oncologica **54**(2), 283-297 (2015).

Introduction

Intensity modulated radiation therapy (IMRT) is particularly useful for treating irregularly shaped planning target volumes (PTVs) whilst minimizing doses to organs-at-risk (OARs). Volumetric modulated arc therapy (VMAT) is a form of rotational IMRT that typically permits shorter delivery times and requires less monitor units (MU) compared to static IMRT beams. Treatment planning aims to create plans with an optimal trade-off between PTV dose coverage and OAR sparing[1–3]. In the search for better VMAT plans there are many parameters that the planner can modify, including the number of arcs. We previously demonstrated that RapidArc® (Varian Medical Systems, Palo Alto, USA) with two arcs was associated with improved plan quality over a single arc and increased agreement between calculated and measured doses, while the beam-on time remained less than 3 minutes[4]. Since the RapidArc optimization algorithm tries to maximize gantry speed we hypothesized that using more than two arcs, which would allow for a longer delivery time and more opportunity for modulation, might translate into further gains in plan quality. We therefore investigated the relationship between the number of arcs, PTV dose homogeneity and OAR sparing.

Materials and Methods

General considerations

Planning CT scans from 10 patients with locally advanced head and neck cancer previously treated using RapidArc were selected for this retrospective planning study. The head and neck location was selected because it is a challenging scenario for treatment planning. Plans were created using the current institutional simultaneous integrated boost (SIB) technique described in detail previously[5]. In brief, this aims to deliver 54.25Gy / 70Gy to the elective / boost PTV (PTV_E / PTV_B) in 35 fractions. PTV_B consisted of the gross tumor volume and biopsy-proven positive lymph nodes with a 5mm margin for microscopic disease (edited for anatomical boundaries) and a 4-5mm PTV margin. PTV_E consisted of elective nodal regions, with a 4-5mm PTV margin, minus a 5mm transition zone (PTV_T), which was created to facilitate a dose fall-off between PTV_B and PTV_E.

OARs with only maximum dose objectives taken into account during optimization included the brainstem, spinal cord, and their planning-at-risk volumes (PRVs) after a 3mm expansion. OARs for which the aim was to reduce the mean dose included the salivary structures (the ipsilateral and contralateral parotid and submandibular glands) and swallowing structures (the upper oesophageal sphincter, upper and lower larynx, the superior, medial and inferior pharyngeal constrictor muscle, cricopharyngeal muscle and

esophagus). The treating clinician may have decided not to spare individual OARs, for example in the case of substantial overlap with the PTVs. Composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) structures (consisting of individually optimized OARs) were created for OAR dose reporting. Data on disease site, stage, PTV and OAR volumes is reported for all patients in Table 1.

Table 1. The disease site, stage and volumes of the elective, transition and boost planning target volumes (PTVs) and the volumes of the composite salivary and swallowing structures for all patients.

Patient Number	Disease Site	Stage	PTV_E [a] (cm^3)	PTV_T [a] (cm^3)	PTV_B [a] (cm^3)	$Comp_{sal}$ [b] (cm^3)	$Comp_{swal}$ [b] (cm^3)
1	Oropharynx	T2N0	346.6	70.2	28.3	41.7	17.3
2	Oropharynx	T2N2b	457.9	34.4	289.0	70.4	18.5
3	Oropharynx	T2N2b	441.7	48.3	188.6	82.3	16.5
4	Nasopharynx	T2N0	538.1	33.9	123.5	105.4	31.7
5	Oropharynx	T2N2a	592.6	104.6	175.6	90.1	12.4
6	Supraglottic larynx	T3N2b	358.5	80.1	143.4	42.4	9.9
7	Oropharynx	T4bN0	422.7	58.6	117.7	86.8	49.6
8	Oropharynx	T3N2c	433.8	103.0	313.0	74.0	17.3
9	Supraglottic larynx	T2N2c	428.7	93.7	243.3	66.7	7.1
10	Oropharynx	TxN2a	240.6	43.4	94.5	42.4	25.0

[a] Elective, transition and boost planning target volumes (PTVs)

[b] Composite salivary and swallowing structures

Treatment planning

Coplanar, 6MV RapidArc plans were created for all 10 patients using two, four, six and eight full arcs optimized using the progressive resolution optimizer (PRO, version 10.0.28; Eclipse treatment planning Varian Medical Systems, Palo Alto, USA). Collimator angles were chosen with a 5° difference between individual arcs. Volume dose was calculated using the Anisotropic Analytical Algorithm (AAA, version 10.0.28) with a 2.5mm dose grid. After a first optimization and dose calculation, a 'continue previous optimization' (CPO) was performed to improve PTV dose homogeneity[5]. All plans were created following our institutional approach to interactive planning[6,7]. In brief, the spinal cord, brainstem and their PRVs were each constrained well below their respective tolerance levels using single maximum dose objectives (optimization priority 120). Because shoulder position can vary slightly during and between fractions, direct irradiation through them was limited (using a maximum dose objective) to increase the robustness of the plan. To minimize the mean dose to the individual salivary glands and swallowing muscles, 4-5 objectives with typical

optimization priorities of 80-90 were used for each structure. The dose-volume location of these objectives was adjusted regularly during the optimization process, aiming to keep them at a constant distance from the OAR dose-volume histogram line that is displayed during optimization. The boost and elective PTVs had minimum and maximum dose objectives placed respectively 1Gy lower and higher than the prescription dose, with PTV optimization priorities (POP) of 150 (POP_{150}).

Plans were evaluated using institutional guidelines. The goal was that at least 99% of the boost volume should be covered by 95% of the prescribed dose (V95% ≥ 99%) and no more than 4% should receive a dose greater than 107% (V107% ≤ 4%). The goals for PTV_E dose homogeneity were V95% ≥ 98% and V107% ≤ 15%.

Study endpoints

Dose inhomogeneity (IH) was calculated for both PTV_B (IH_B) and PTV_E (IH_E) in all plans using the equation IH = (100% − V95%) + V107%. This metric consists of the PTV dose values that are analyzed clinically in our institute (V95% and V107%) and is consistent with ICRU reports 50 and 62. To allow for a broader comparison of our results, the homogeneity index (HI) was also calculated for both the boost (HI_B) and elective (HI_E) PTVs using the equation suggested by ICRU 83: HI = 100% × (D2% − D98%)/D50%. Mean doses to $comp_{sal}$ and $comp_{swal}$ were averaged to provide a composite mean OAR (OAR_{comp}) dose. To investigate whether additional arcs increase dose deposition in the remaining body, the body minus PTV volume receiving a dose ≥5Gy (V5Gy), ≥30Gy (V30Gy) and ≥50Gy (V50Gy) was determined. An ANOVA test with post-hoc Bonferroni analysis was performed (using IBM SPSS, Armonk, NY) to investigate which of the analyzed parameters improved significantly using more arcs. Finally, we reported the number of monitor units (MUs) as a surrogate for plan modulation.

To investigate the effect of using more arcs for different PTV optimization priorities, one to eight arc plans were made for patients 1 to 3 at three POP levels: POP_{130}, POP_{150} and POP_{170}. The boost and elective PTVs used the same POP value.

Dosimetric validation was performed for patients 1 to 3 by measuring the dose distributions of the two, four, six and eight arc plans made with a PTV optimization priority of 150 (POP_{150}) on a MatriXX ionisation array (IBA Dosimetry, Schwarzenbruck, Germany) placed in a box phantom, and comparing these with calculated doses on the CT scan of the MatriXX / box phantom (OmniPro I'mRT software, IBA Dosimetry) using gamma evaluation with a 2mm distance to agreement and 3% dose difference. To investigate the influence of calculation resolution on gamma analysis for plans with more MUs[8] the evaluation of the four, six and eight arc plans was repeated for dose calculations using a 1.0mm calculation grid.

Delivery time is expected to increase linearly with the number of arcs because RapidArc attempts to generate plans at maximum gantry speed, and gantry speed is only reduced if the dose rate is at maximum. As a result, the delivery time per arc does not vary as long as the gantry is rotating at maximum speed. Therefore, delivery time, from the start of the first arc until the end of the last arc, was only measured for two, four, six and eight-arc plans of one patient at POP_{150} on the TrueBeam (Varian Medical Systems).

Results

Maximum dose constraints to the brainstem and spinal cord were met in all plans. Increasing the number of arcs increased PTV dose homogeneity, OAR sparing and MU count (Table 2). The gains in OAR sparing were similar for the individual salivary structures (supporting the use of composite salivary structures and OAR_{comp} for OAR dose reporting). Except for the body – PTV doses, IH_B and PTV_B V95% and V107%, all parameters improved significantly when four instead of two arcs were used. The magnitude of the gains decreased going from four to six and four to eight arcs, and no statistically significant differences were obtained when increasing from six to eight arcs.

Figure 1 shows the effect of increasing the number of arcs on IH_B, IH_E, $comp_{sal}$ and $comp_{swal}$ for all individual patients. Although the initial values and subsequent gains in these parameters vary, all plans benefitted from additional arcs either by lowering IH_B / IH_E, OAR doses, or both. When a homogeneous PTV dose was already obtained with two arcs, more arcs could still improve OAR sparing (e.g. in patient 1). In contrast, in some cases where there was only a minor reduction in OAR dose, bigger gains in IH_B and IH_E were possible (e.g. in patient 9). For some patients with large PTV_B volumes (such as patients 2, 8 and 9), it was hard to obtain adequate PTV_B coverage using two arcs, resulting in high IH_B values. Adding more arcs in these cases resulted in improvements in IH_B. HI_B and HI_E were also consistently lower using additional arcs. V5Gy and V30Gy both tended to increase using more arcs, although this increase was no more than 1.7% and 0.4%, respectively, and therefore unlikely to be clinically relevant. V50Gy decreased by less than 0.02% when using more arcs.

Table 2. The effect of arc number on boost and elective planning target volume (PTV) dose homogeneity and organ-at-risk (OAR) doses.

Number of Arcs	Two	Four	Six	Eight
	Boost PTV (%)			
V95%[a]	99.3 ± 0.7	99.6 ± 0.5	99.7 ± 0.3	99.7 ± 0.4
V107%[a]	0.9 ±1.2	0.4 ± 0.9	0.2 ± 0.5	0.1 ± 0.3
IH[b]	1.6 ± 1.6	0.8 ± 1.0	0.5 ± 0.6	0.4 ± 0.5
HI[c]	9.2 ± 1.5	8.4 ± 1.4 *	7.7 ± 1.1 * †	7.5 ± 1.1 * †
	Elective PTV (%)			
V95%[a]	98.6 ± 0.5	99.0 ± 0.4 *	99.2 ± 0.4 * †	99.2 ± 0.4 * †
V107%[a]	10.8 ± 3.4	8.1 ± 3.0 *	6.5 ± 2.4 * †	5.5 ± 1.9 * †
IH[b]	12.2 ± 3.2	9.1 ± 3.1 *	7.4 ± 2.5 * †	6.2 ± 1.9 * †
HI[c]	14.0 ± 1.2	12.7 ± 1.6 *	11.8 ± 1.3 * †	11.4 ± 1.0 * †
	Mean dose (Gy)			
CL Parotid (n=10)	24.1 ± 7.5	22.2 ± 7.4 *	21.4 ± 6.9 *	21.0 ± 6.8 *
IL Parotid (n=10)	30.5 ± 6.7	28.2 ± 6.2 *	26.8 ± 5.6 *	26.4 ± 5.6 *
CL SMG[d] (n=9)	38.2 ± 7.2	35.6 ± 7.7 *	34.3 ± 7.3 *	34.5 ± 7.3 * †
$Comp_{sal}$[e]	28.2 ± 6.5	26.0 ± 6.3 *	24.9 ± 5.8 *	24.5 ± 5.6 *
$Comp_{swal}$[e]	30.0 ± 8.0	27.5 ± 8.1 *	26.6 ± 8.5 * †	26.0 ± 8.3 * †
OAR_{comp} dose[f]	29.1 ± 4.0	26.8 ± 4.5 *	25.7 ± 4.8 * †	25.3 ± 4.7 * †
	Body - PTV (cm^3)			
V5Gy[g]	5258 ± 1355	5310 ± 1357	5318 ± 1347 *	5346 ± 1345 * †
V30Gy[g]	1882 ± 519	1917 ± 506	1943 ± 512 *	1969 ± 514 * †
V50Gy[g]	338 ± 130	335 ± 124	325 ± 117	332 ± 117
Monitor Units	473 ± 34	543 ± 40	595 ± 48	619 ± 39

* / † Significant differences compared to the two / four arc plans

[a] V95% / V107% PTVs receiving ≥95% / ≤107% prescribed dose

[b] PTV dose inhomogeneity (100% – PTV 95% + PTV V107%)

[c] PTV homogeneity index (100% × [D2% – D98%] / D50%)

[d] Contralateral submandibular gland

[e] Composite salivary and swallowing structures

[f] OAR volume receiving average of mean doses to compsal and compswal

[g] Body-PTV volumes receiving 5Gy / 30Gy / 50Gy

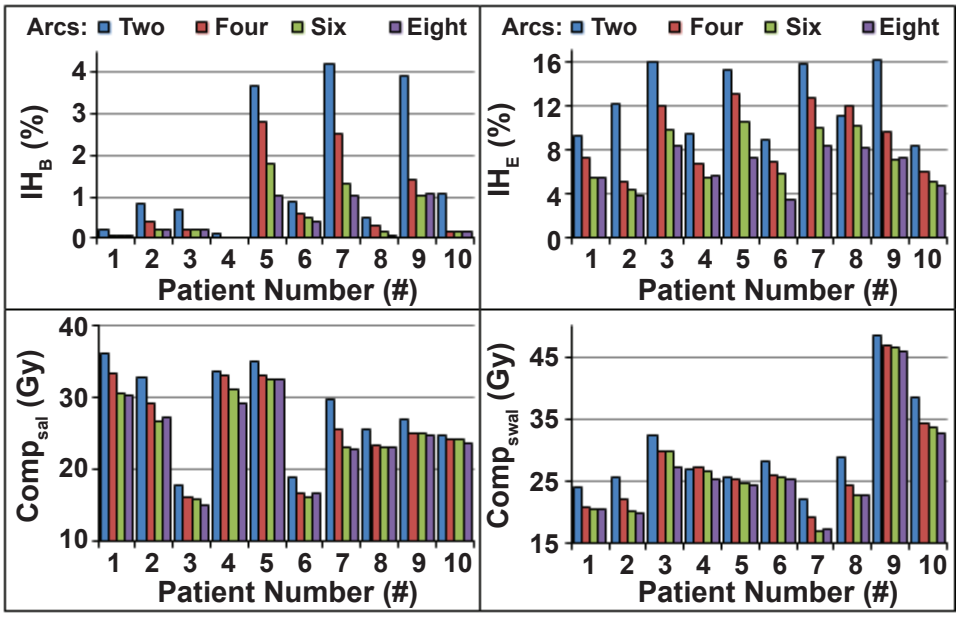


Figure 1. Histograms showing the effect of using two, four, six or eight arcs on boost and elective dose inhomogeneity and mean dose to the composite salivary and swallowing structures for all patients.

The results of using one to eight arcs on IH_B (left graph), IH_E (right graph) and OAR_{comp} dose is shown in Figure 2 for three levels of POP and for patients 1-3. The colors of the datapoints indicate the number of arcs that were used. This Figure clearly demonstrates that both OAR dose and PTV dose inhomogeneity tend to decrease with increasing number of arcs. The relative gains however differed depending on the chosen POP value. With the lowest PTV dose inhomogeneity values, adding additional arcs at POP_{170} primarily improved OAR sparing. In contrast, plans made with POP_{130} resulted in the most OAR sparing and adding additional arcs primarily improved PTV dose homogeneity. The largest gains were achieved when increasing the number of arcs from one to two and two to four. Closer grouping of the four to eight arc plans indicates diminishing absolute returns above four arcs.

Gamma analysis of the POP_{150} plans of patients 1 to 3 showed an average pass rate of 99.6%, 98.3%, 96.3% and 98.1% for the two, four, six and eight arc plans, respectively. In patients 1 and 2, the six and eight arc plans showed the weakest agreement between measured and calculated doses. Recalculation of the four, six and eight arc plans at high-resolution generally resulted in a better pass rate, averaging 99.1%, 99.3% and 98.3%, respectively.

High-resolution calculations did not influence IH_B / IH_E but marginally reduced OAR_{comp} mean dose (0.2-0.5Gy). Delivery times for the two, four, six and eight arc plans of patient 1 were 125, 260, 387 and 529 seconds, respectively.

Discussion

In this systematic investigation of the effect of arc number on VMAT plan quality using RapidArc we found a clear trend for improved OAR sparing and more homogeneous PTV doses going from one to eight arcs in head and neck VMAT plans with a simultaneous integrated boost. The largest absolute gains were found when increasing the number of arcs from two, which is most commonly used in clinical practice, to four. Although the improvements were smaller, further gains were still possible using more than four arcs. Taking into account the available literature regarding normal tissue toxicity[9–12], we believe that the dosimetric gains are sufficient to be clinically relevant. The resulting increase in plan quality comes at the expense of increased delivery time, and a modest increase in MU, both of which we consider clinically acceptable. To put this increased delivery time into context, our four arc plans were delivered in under 4.5 minutes whereas the single arc SmartArc plans in the study of Lechner et al.[3] could take more than 6 minutes. This illustrates that data concerning the use of multiple arcs from other systems, for example SmartArc (Philips Healthcare) or VMAT (Elekta) cannot be extrapolated to RapidArc because of fundamental differences in the treatment planning and delivery system[13–19]. This makes it necessary to perform system-specific treatment planning studies like the present report. Whether these results can be generalized to other anatomical locations or to other versions of the treatment planning system requires further evaluation.

One reason for improved plan quality when using more arcs might be specific characteristics of the RapidArc optimizer. It tends to create a treatment plan in which the gantry rotates at maximum speed, while the dose rate and leaf positions are modulated. The gantry speed is only reduced if the maximum dose rate prevents sufficient delivery of dose. Our results suggest that two arcs delivered at maximum gantry speed cannot achieve sufficient fluence modulation to create an optimal plan. It is therefore possible that RapidArc plans with one or two arcs could be created with similar quality to four or more arcs if the optimizer allowed gantry slowdown.

In conclusion, using the clinical scenario of complex head and neck radiotherapy, we have quantitatively demonstrated increased plan quality when more than two arcs are used. The four arc plans seemed to provide a good trade-off between increased delivery time and improved plan quality. This data also illustrates that gains in complex photon plans are still possible using existing technology.

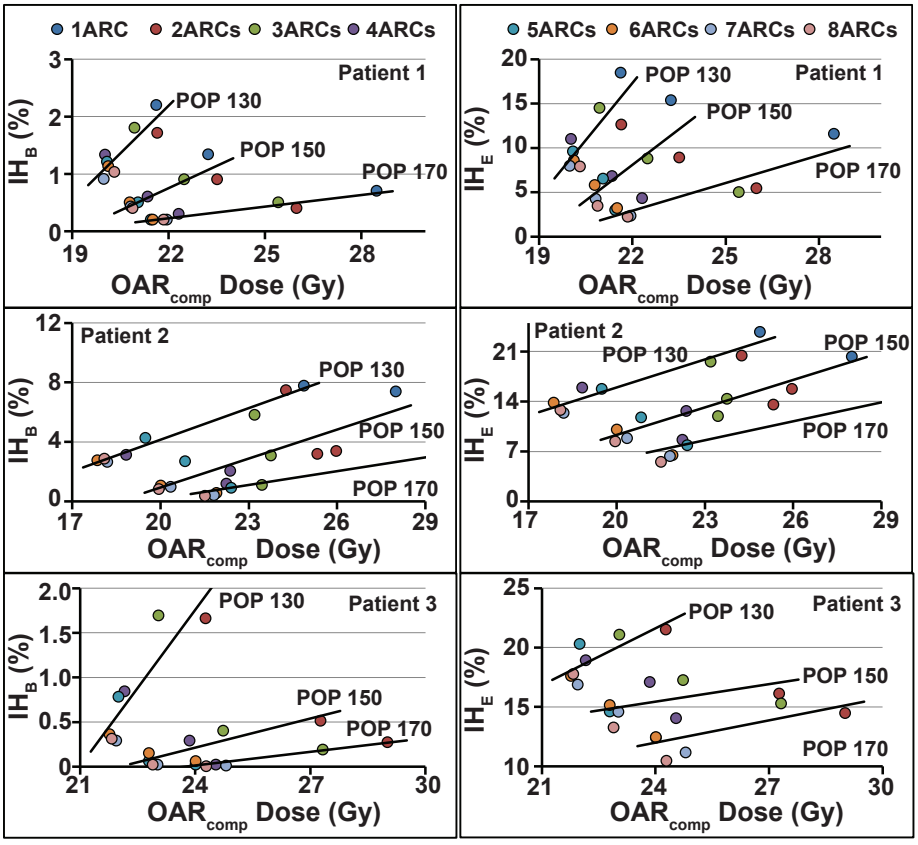


Figure 2. Planning target volume (PTV) dose inhomogeneity for PTV boost (IH_B, left graphs) and PTV elective (IH_E, right graphs) versus composite organ-at-risk (OAR_{comp}) mean dose for patients 1 to 3 and three different PTV optimization priorities (POP). The solid lines represent linear fits through the datapoints of the POP_{130}, POP_{150} and POP_{170} plans. The different colors indicate the number of arcs used per datapoint. The one arc datapoints are omitted in the graph of patient 3 because for this patient it proved impossible to obtain adequate PTV dose coverage and OAR sparing using only a single arc. This resulted in large deviations from the other datapoints thereby hampering analysis of the results.

References

1 R.O. Ottosson, P.E. Engstrom, D. Sjöström, C.F. Behrens, A. Karlsson, T. Knöös, and C. Ceberg, "The feasibility of using Pareto fronts for comparison of treatment planning systems and delivery techniques.," Acta Oncol. **48**(2), 233–7 (2009).

2 D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning.," Med. Phys. **33**(9), 3399–407 (2006).

3 W. Lechner, G. Kragl, and D. Georg, "Evaluation of treatment plan quality of IMRT and VMAT with and without flattening filter using Pareto optimal fronts.," Radiother. Oncol. **109**(3), 437–41 (2013).

4 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–9 (2009).

5 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

6 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

7 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

8 C.L. Ong, J.P. Cuijpers, S. Senan, B.J. Slotman, and W.F.A.R. Verbakel, "Impact of the calculation resolution of AAA for small fields and RapidArc treatment plans.," Med. Phys. **38**(8), 4471–9 (2011).

9 I. Beetz, C. Schilstra, A. van der Schaaf, E.R. van den Heuvel, P. Doornaert, P. van Luijk, A. Vissink, B.F.A.M. van der Laan, C.R. Leemans, H.P. Bijl, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, and J.A. Langendijk, "NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors.," Radiother. Oncol. **105**(1), 101–6 (2012).

10 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F.A.M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

11 T. Dijkema, C.P.J. Raaijmakers, R.K. ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. 78(2), 449–53 (2010).

12 J.O. Deasy, V. Moiseenko, L. Marks, K.S.C. Chao, J. Nam, and A. Eisbruch, "Radiotherapy dose-volume effects on salivary gland function.," Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl), S58–63 (2010).

13 J.L. Bedford, "Treatment planning for volumetric modulated arc therapy," Med. Phys. **36**(11), 5128 (2009).

14 D. Cao, M.K.N. Afghan, J. Ye, F. Chen, and D.M. Shepard, "A generalized inverse planning tool for volumetric-modulated arc therapy.," Phys. Med. Biol. **54**(21), 6725–38 (2009).

15 C. Cameron, "Sweeping-window arc therapy: an implementation of rotational IMRT with automatic beam-weight calculation.," Phys. Med. Biol. **50**(18), 4317–36 (2005).

16 A. Chi, P. Ma, G. Fu, G. Hobbs, J.S. Welsh, N.P. Nguyen, S.Y. Jang, J. Dai, J. Jin, and R. Komaki, "Critical structure sparing in stereotactic ablative radiotherapy for central lung lesions: helical tomotherapy vs. volumetric modulated arc therapy.," PLoS One **8**(4), e59729 (2013).

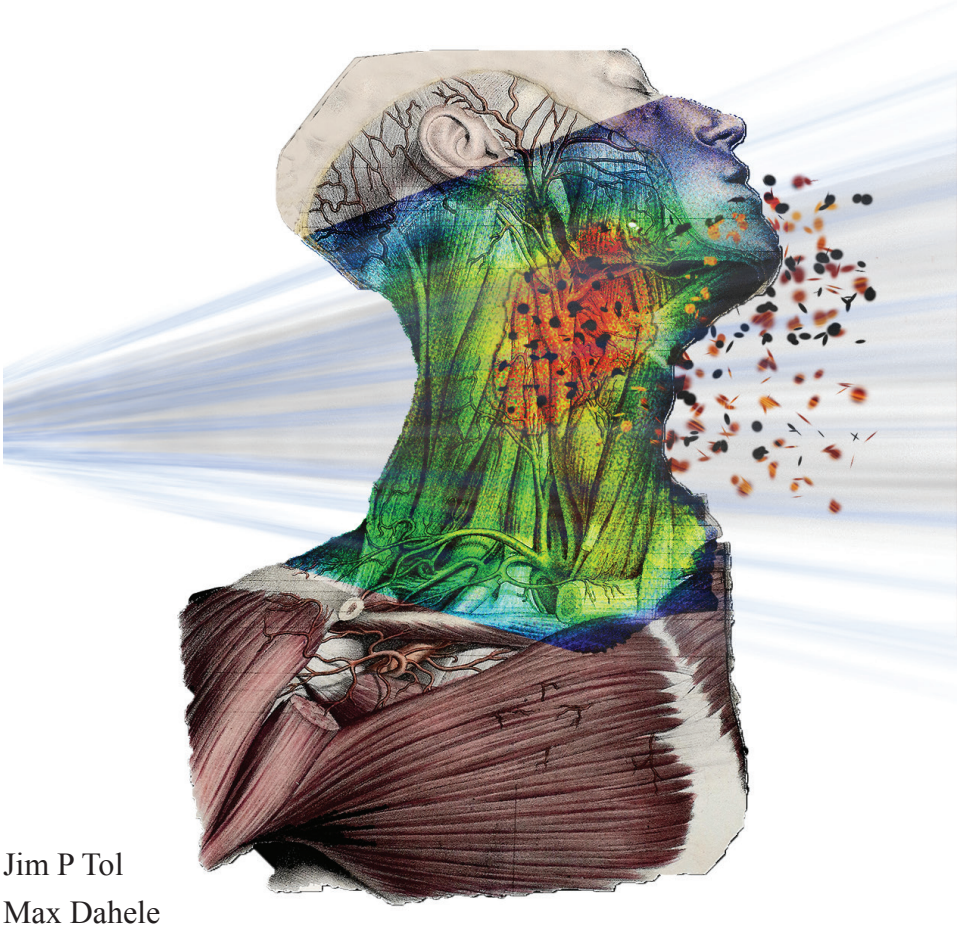
17 M. Guckenberger, A. Richter, T. Krieger, J. Wilbert, K. Baier, and M. Flentje, "Is a single arc sufficient in volumetric-modulated arc therapy (VMAT) for complex-shaped target volumes?," Radiother. Oncol. **93**(2), 259–65 (2009).

18 E. Vanetti, G. Nicolini, J. Nord, J. Peltola, A. Clivio, A. Fogliata, and L. Cozzi, "On the role of the optimization algorithm of RapidArc(®) volumetric modulated arc therapy on plan quality and efficiency.," Med. Phys. **38**(11), 5844–56 (2011).

19 K. Bzdusek, H. Friberger, K. Eriksson, B. Hårdemark, D. Robinson, and M. Kaus, "Development and evaluation of an efficient approach to volumetric arc therapy planning," Med. Phys. **36**(6), 2328 (2009).

Chapter 3

Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: an exponential trade-off with target volume dose homogeneity



Jim P Tol
Max Dahele
Patricia Doornaert
Ben J Slotman
Wilko FAR Verbakel

Medical Physics **41**(2), 021722 (2014).

Abstract

Purpose

Conventional radiotherapy typically aims for a homogenous dose in the planning target volume (PTV) while sparing organs-at-risk (OAR). We quantified and characterized the trade-off between PTV dose inhomogeneity (IH) and OAR sparing in complex head and neck volumetric modulated arc therapy (VMAT) plans.

Materials and Methods

13 simultaneous integrated boost plans were created per patient, for 10 patients. PTV boost / elective (PTV_B / PTV_E) optimization priorities were systematically increased. IH_B and IH_E, defined as (100% – PTV V95%) + PTV V107% was evaluated against the average of the mean doses to the combined composite swallowing and salivary organs (OAR_{comp}). To investigate the influence of OAR size and position with respect to PTV_B and PTV_E, the OAR doses were evaluated against a modified Euclidean distance (DM_B / DM_E) between the OARs and PTVs.

Results

Although the achievable OAR_{comp} dose for a given level of PTV IH differed between patients, excellent logarithmic fits described the relationship between OAR_{comp} dose and IH_B / IH_E in all patients (mean R^2 of 0.98 / 0.97). Allowing for an increase in average IH_B and IH_E over a clinically acceptable range, e.g. from 0.4±0.5% to 2.0±2.0% and 6.9±2.8% to 14.8±2.7%, respectively, corresponded to a decrease in average dose to the composite salivary and swallowing structures from 30.3±6.5Gy to 23.6±4.7Gy and 32.5±8.3Gy to 26.8±9.3Gy. The increase in PTV_E IH was mainly accounted for by an increase in V107, by on average 5.9%, rather than a reduction in V95, which was on average only 2%. A linear correlation was found between DM_E and the OAR dose to composite swallowing structures and contralateral parotid and submandibular glands (R^2 values of 0.83, 0.88 and 0.95, respectively). Only the mean ipsilateral parotid dose correlated with DM_B (R^2=0.87).

Conclusions

OAR sparing is highly dependent on the permitted PTV IH. PTV_E IH substantially influences OAR doses. These results are relevant for clinical practice and for future automated treatment planning strategies.

Introduction

High-quality radiotherapy treatment plans typically aim to cover the planning target volume (PTV) with a homogeneous dose whilst minimizing dose to nearby organs-at-risk (OARs). This is particularly hard to achieve when multiple PTVs and OARs are in close proximity to each other and the resulting plans can therefore vary substantially between treatment planners and radiotherapy centers[1,2]. Although automated and knowledge-based planning strategies have been proposed as a means of reducing such variation and promising results have already been obtained[3–8], considerable variation exists in protocols for acceptable PTV dose coverage and homogeneity. Several previous studies[9–12], which used plans with only a limited number of OARs, have shown that OAR sparing is influenced by the level of PTV dose homogeneity. However, to the best of our knowledge, a systematic, quantitative analysis has not been carried out to highlight the extent to which emphasizing PTV dose homogeneity may compromise OAR sparing in more complex clinical head and neck volumetric modulated arc therapy (VMAT) plans. Increased potential for OAR sparing has been the key driver in the use of advanced technologies like VMAT in head and neck cancer radiotherapy. Understanding PTV-OAR trade-offs is therefore important to be able to generate plans that provide optimal OAR sparing at a given level of PTV dose homogeneity. We studied these trade-offs using clinically relevant ranges of target volume dose inhomogeneity. We also investigated the relationship between OAR dose and the elective / boost PTV-OAR distance[6,13,14] for plans at a given level of PTV dose inhomogeneity. This work was considered a necessary step in the development of (semi)-automated treatment planning strategies for high-quality head and neck cancer VMAT.

Materials and Methods

Planning CT scans from 10 head and neck cancer patients previously treated using RapidArc were arbitrarily selected for this analysis. Table 1 shows tumor site and stage for every patient. All plans were created using the institutional simultaneous integrated boost (SIB) technique delivering 54.25Gy to the elective PTV (PTV_E) and 70Gy to the boost PTV (PTV_B) in 35 fractions of 1.55 and 2Gy, respectively. The PTV_B region consisted of the gross tumor volume (GTV, typically delineated using CT and MRI fusion) and biopsy-proven positive lymph nodes, with a 5mm margin for microscopic disease (edited for anatomical boundaries) and a 4-5mm PTV margin. PTV_E consisted of elective nodal regions, with a 4-5mm PTV margin, minus a 5mm transition zone (PTV_T) between PTV_B and PTV_E, created to facilitate a steep dose falloff between the PTVs. For optimization and dose reporting purposes, a 'virtual' build-up region[15] was routinely created where the PTV

extended outside the body by expanding the body 5mm beyond the PTV. All dosimetric parameters are reported using this structure.

Constrained OAR were as they appeared in the clinical plan. These included the brainstem and spinal cord (and their planning at risk volumes [PRVs] after 3mm expansion), salivary and swallowing structures. The salivary structures could consist of the ipsilateral and contralateral parotid and submandibular glands. The swallowing structures could consist of the upper oesophageal sphincter, upper and lower parts of the larynx, the superior, medial and inferior pharyngeal constrictor muscle, the cricopharyngeal muscle and esophagus[16]. In general, when there was considerable overlap between PTV and OAR, the decision whether to try and spare the OAR was made by the clinician.

Plan optimization

Coplanar RapidArc plans were made in the Eclipse treatment planning system (TPS), version 10.0.28, (Varian Medical Systems, Palo Alto, USA) using two full 6MV arcs (rotating clockwise and counter-clockwise).Volume dose was calculated using the Anisotropic Analytical Algorithm (AAA) with a 2.5mm dose grid. After a first optimization and dose calculation, a 'continue previous optimization' (CPO) was performed. This CPO uses the AAA calculated dose as a starting point for an additional optimization in order to compensate for differences between the simplified dose calculation algorithm used during optimization and final dose calculation with AAA.

All plans were interactively optimized following our previously described institutional planning protocol[17–19]. The spinal cord, brainstem and their PRVs were constrained well below their respective dose tolerance levels using a single maximum dose objective (priority 120). These objectives were never interactively adjusted because the dose tolerance levels were never exceeded. The shoulders were also routinely constrained to limit direct irradiation through them, since the shoulder position can vary slightly between fractions. To minimize the mean dose to the individual salivary and swallowing structures, 4-5 objectives with optimization priorities of 80-90 were used. During the optimization process, these objectives were adjusted interactively, aiming to keep them at fixed distance from their dose-volume histogram (DVH) line that is displayed during the optimization process. If there was overlap between the OAR and PTV, optimization was performed on OAR-PTV (the part of the OAR without PTV overlap) instead. Dose reporting is done for the entire OAR structure. The PTVs had minimum and maximum dose objectives respectively 1Gy lower and higher than the prescription dose.

Our institutional guidelines for PTV_B dose coverage and homogeneity in this clinical scenario aim to deliver 95% of the prescribed dose to at least 99% of the boost volume (V95 ≥ 99%) and no more than 4% of this volume to receive a dose greater than 107% (V107 ≤ 4%). This is typically achieved by normalizing the treatment plan such that the prescribed dose is delivered to approximately 70-80% of PTV_B and the mean PTV_B dose falls within 101-103% of the prescribed dose. Acceptable PTV_E dose homogeneity is less strict and aims for V95 ≥ 98% and V107 ≤ 15%. It is worth noting that a large variety of clinical plan acceptance criteria exist between institutes and studies. Some view the prescription dose as the minimum dose to be received by a specific volume of the PTV, while others view the prescription dose as the mean or median dose that the PTV should receive. This point is expanded upon in the discussion.

For each of the 10 patients, 13 treatment plans were created using the interactive optimization method described above. Priorities of the OAR constraints were kept constant in each optimization whereas the PTV optimization priorities (POP) were systematically increased by a value of 10, resulting in a PTV priority range of 80-200 (POP_{80}-POP_{200}) in 13 increments. The higher the POP value, the more effort is spent by the optimization algorithm to achieve homogeneous PTV doses. Because the OAR optimization priorities are kept constant, this allowed us to investigate the trade-off between PTV dose homogeneity and OAR sparing. All plans were normalized such that they delivered the same mean PTV_B dose as the original clinical plan. It was therefore not our goal that all plans met the institutional goals for V95 or V107. In this way we have been able to test our hypothesis that a slight reduction in PTV coverage and / or a larger volume of higher doses (i.e. more dose heterogeneity, but same mean dose), could lead to significant gains in OAR doses. In all plans, a measure of the dose inhomogeneity was calculated for both PTV_B (IH_B) and PTV_E (IH_E) using the equation IH = (100% – V95%) + V107%. Composite swallowing and salivary structures were created for OAR dose reporting. The mean doses to these structures were averaged to a composite OAR (OAR_{comp}) dose. For each patient, IH_B and IH_E were plotted against OAR_{comp} dose and exponential fits were generated using the curve-fitting tool of MATLAB (The MathWorks Inc). Each OAR_{comp} dose value corresponds to one pair of IH_B and IH_E values.

To investigate whether increased sparing of swallowing and salivary structures leads to additional dose deposition in the remaining body, the mean dose in an additional structure (Body – [OAR + PTV]) consisting of Body (in longitudinal direction extending 2cm outside the last slice with PTV) minus the PTVs and OAR_{comp} was evaluated. In addition, the volume of the Body – (OAR + PTV) structure receiving a dose >50Gy (V50Gy) and >60Gy (V60Gy) was reported. Since performing a CPO can reduce the PTV dose inhomogeneity

considerably, the combination of CPO and a lower POP can be an important tool to improve OAR sparing. The effect of CPO calculations on IH_E, IH_B and OAR_{comp} dose was therefore investigated for all patients.

OAR to PTV distance

For automated planning strategies, it can be important to be able to predict the maximum achievable amount of OAR sparing at a prescribed target dose coverage and homogeneity for a new patient. Using a method introduced by Kazhdan et al.[6,20], Yuan et al.[21] showed previously that achievable OAR dose is related to a distance measure (DM) which is based on the Euclidean Distance Function. We therefore investigated the relationship between PTV dose inhomogeneity and OAR sparing and the dependence of OAR dose on the *DM* between OAR and PTV. This is demonstrated for one IH_B value, but can be expanded similarly to other IH_B / IH_E values.

For every OAR voxel (*i*), the smallest distance to the PTV boundary was determined, using 2.5mm voxel size. If *i* was located within the PTV, a negative value was assigned to DM_i. Because OAR voxels located cranially or caudally from CT-slices containing PTV receive mostly scattered dose, the effective distance measure for those voxels was increased linearly by 1 + 0.1 × (craniocaudal voxel distance to the PTV). *DM* is obtained by averaging DM_i over all OAR voxels. DM_B and DM_E represent the distance measures of the OAR to PTV_B and PTV_E, respectively.

We investigated the relationship between DM_B / DM_E and the mean dose to the individual salivary glands and composite swallowing structures. Mean dose to these structures was obtained by creating exponential fits between IH_B and dose to these organs for all patients and evaluating these at 5% IH_B. The salivary glands were evaluated separately because of the large differences in distance to the PTVs (especially PTV_B). The ipsilateral submandibular gland generally showed large overlap with the PTVs and was therefore not attempted to be spared in 6 out of 10 patients. Since the individual swallowing organs are typically small and located closely to each other, we chose to investigate the relationship between DM_B / DM_E and mean dose to the composite swallowing structure.

Results

PTV volumes for all patients are shown in Table 1 along with the size of the composite salivary and swallowing structures. The large variation between patients in size of the composite OAR is caused by the exclusion of different individual OARs that could not be spared because they overlapped considerably with the PTV.

Table 2 summarizes the plan results, averaged over all patients. Clinically acceptable brainstem and spinal cord maximum doses were achieved in all plans (38.2±9.7Gy and 35.3±7.8Gy on average, respectively) regardless of POP. A clear trend was such that a decrease in POP lead to increased IH_B / IH_E and decreased OAR_{comp} doses. Comparing POP_{200} with POP_{120} plans (mean V95 of both PTVs > 95%), an increase of IH_B and IH_E from 0.4±0.5% to 4.7±3.7% and 6.9±2.8% to 21.7±3.7%, respectively, corresponded to a decrease in dose to the composite salivary and swallowing structures from 30.3±6.5Gy to 22.4±5.1Gy and 32.5±8.3Gy to 24.3±8.7Gy. Plans in the high POP range show that considerable gains in OAR sparing can be achieved within a smaller range of IH_B / IH_E values as well. For example, comparing the POP_{200} and POP_{170} plans illustrates that a relatively small increase in IH_B and IH_E from 0.4% to 1.1% and 6.9% to 10.8%, respectively, allows for a decrease in OAR_{comp} dose from 31.0Gy to 26.7Gy. This was largely accounted for by an increase in PTV_B and PTV_E V107, while PTV coverage with 95% of the prescription dose was affected to a lesser extent: PTV_B and PTV_E V95 decreased from 99.8% to 99.6% and 99.4% to 98.7%, respectively. These results demonstrate that a small relaxation of PTV V107 may contribute substantially to OAR sparing while the minimum PTV coverage is little affected.

Figure 1 shows IH_B (lower-) and IH_E (upper y-axis) plotted against OAR_{comp} dose (x-axis) for every patient. The exponential fits through the individual data points are represented by the solid, dashed and dotted lines, with mean R^2 correlation coefficients of 0.98±0.01 and 0.97±0.03 for IH_B and IH_E respectively, indicating a similar trade-off between IH_B / IH_E and OAR_{comp} dose in every patient. The slopes show for all patients that at large IH_B (low POP), further increasing PTV dose inhomogeneity leads to only a marginal decrease in OAR_{comp} dose, whereas for plans with a very homogeneous PTV dose (low IH_B, high POP), a small increase of IH_B may lead to a relatively large reduction in OAR_{comp} dose. For the individual salivary and swallowing structures, comparable exponential fits were obtained between IH_B / IH_E and their mean doses. This supports the use of composite salivary and swallowing structures and OAR_{comp} for OAR dose reporting.

In order to visualize the increased OAR sparing for lower POP plans, Figure 2 shows dose profiles through the OAR and PTV for plans with different POP parameters. For patients 7 and 8, the dose profile of the POP_{150} plans (dotted line) at the level of the parotid glands (orange and blue contours) and the cricopharyngeal muscle (orange contour) are compared with dose profiles of the POP_{130} (solid line) and POP_{200} (dashed line) plans. For patient 7, dose to the cricopharyngeal muscle varies substantially with the level of POP, whereas this is not the case in patient 8. This may be because the boost PTV is only present in the chosen plane in patient 8, thereby limiting the potential for OAR sparing at this level. At the level

of the parotid glands the trends seem similar for both patients. The POP_{200} plans provide the least fluctuation of dose within the PTVs whereas the dose outside these is consistently higher. In contrast, the POP_{130} plans exhibit the steepest dose fall-off outside the PTVs, resulting in lower OAR doses and the highest fluctuation of dose within the PTVs.

Figure 3 demonstrates for every patient at POP_{150} that CPO calculations reduce IH_B and IH_E, whereas OAR_{comp} dose is less affected. Averaged over all patients, CPO calculations reduced IH_B / IH_E by 1.8±1.1% / 3.1±3.8%.

Averaged over all patients, mean dose to Body – (OAR + PTV) reduced with decreasing POP from 19.8±1.6Gy (POP_{200}) to 18.9±1.7Gy (POP_{80}). This was also apparent from the V50Gy / V60Gy determined for all patients. On average, these values decreased from 5.2% / 0.6% to 3.8% / 0.5% comparing the most and least homogeneous plans respectively, indicating that increased sparing of delineated OARs does not result in an increased high dose deposition in the remainder of the body.

The upper graph of Figure 4 shows the trade-off curves (R^2=0.96±0.03) between IH_B and mean contralateral parotid gland dose. The mean contralateral parotid gland dose corresponding to 5% IH_B (horizontal dotted line) of every curve is plotted against DM_B (circles) and DM_E (plusses) in the lower graph. Good correlation (R^2=0.88) of the linear fit demonstrates that for these plans the achievable contralateral parotid gland mean dose at 5% IH_B can be estimated by DM_E. The relation with DM_B however is very poor (R^2=0.01) indicating that distance of the contralateral parotid gland to the boost PTV is no indicator of achievable dose.

Figure 5 shows the mean dose at 5% IH_B (x-axis) to the ipsilateral parotid gland, contralateral submandibular gland and composite swallowing structures plotted against DM_B (circles, y-axis) and DM_E (plusses, y-axis) for all patients, along with solid linear fits. The mean doses to the composite swallowing structures and submandibular glands correlated well with DM_E, with R^2 values of 0.83 and 0.95, respectively. The mean dose to the ipsilateral parotid gland showed good correlation with DM_B (R^2=0.87). These results demonstrate that the mean dose of various OAR at a chosen level of IH_B / IH_E is clearly influenced by the PTV-OAR geometry (in particular, the distance to the nearest PTV) and can be estimated by DM_E / DM_B.

Table 1. The disease site, stage and volumes of the elective, transition and boost planning target volumes (PTV_E, PTV_T, PTV_B, respectively) and the volumes of the composite salivary ($comp_{sal}$) and swallowing structures ($comp_{swal}$) for all patients.

Patient Number	Disease Site	Stage	PTV_E (cm^3)	PTV_T (cm^3)	PTV_B (cm^3)	$comp_{sal}$ (cm^3)	$comp_{swal}$ (cm^3)
1	Oropharynx	T2N0	346.6	70.2	28.3	41.7	17.3
2	Oropharynx	TxN2a	240.6	43.4	94.5	42.4	25.0
3	Nasopharynx	T2N0	538.1	33.9	123.5	105.4	31.7
4	Supraglottic larynx	T3N2b	358.5	80.1	143.4	42.4	9.9
5	Oropharynx	T3N2c	433.8	103.0	313.0	74.0	17.3
6	Oropharynx	T2N2a	592.6	104.6	175.6	90.1	12.4
7	Oropharynx	T2N2b	441.7	48.3	188.6	82.3	16.5
8	Oropharynx	T2N2b	457.9	34.4	289.0	70.4	18.5
9	Oropharynx	T4bN0	422.7	58.6	117.7	86.8	49.6
10	Supraglottic larynx	T2N2c	428.7	93.7	243.3	66.7	7.1

Table 2. Boost and elective planning target volume (PTV) dose homogeneity and mean dose to the composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) structures. OAR_{comp} represents the average of the mean doses to $comp_{sal}$ and $comp_{swal}$.

POP[a]	PTV_B[b] V95 (%)	PTV_E[b] V95 (%)	PTV_B V107 (%)	PTV_E V107 (%)	IH_B[c] (%)	IH_E[c] (%)	$comp_{sal}$ (Gy)	$comp_{swal}$ (Gy)	OAR_{comp}[d] (Gy)
80	92.2 ± 2.9	83.1 ± 7.0	15.6 ± 6.8	23.3 ± 5.9	23.3 ± 8.7	40.2 ± 5.7	18.6 ± 4.4	20.7 ± 8.0	20.4 ± 5.7
90	94.4 ± 2.4	87.5 ± 4.7	11.7 ± 6.3	22.3 ± 5.4	17.4 ± 8.1	34.8 ± 5.0	19.8 ± 4.4	21.4 ± 8.5	21.3 ± 5.8
100	96.2 ± 2.0	92.3 ± 2.9	7.5 ± 5.4	21.4 ± 6.6	11.3 ± 6.9	29.1 ± 5.8	20.6 ± 4.4	22.2 ± 8.6	22.3 ± 6.2
110	97.1 ± 1.4	93.9 ± 2.3	4.9 ± 4.3	19.4 ± 6.0	7.9 ± 5.3	25.5 ± 5.4	21.1 ± 4.5	23.2 ± 8.7	23.0 ± 6.2
120	98.2 ± 1.1	95.7 ± 1.7	2.9 ± 2.9	17.4 ± 3.9	4.7 ± 3.7	21.7 ± 3.7	22.4 ± 5.1	24.3 ± 8.7	24.3 ± 5.3
130	98.7 ± 0.7	96.8 ± 1.0	2.0 ± 2.1	14.9 ± 3.4	3.3 ± 2.5	18.1 ± 3.4	23.6 ± 5.1	25.6 ± 8.8	25.4 ± 6.4
140	99.1 ± 0.6	97.4 ± 0.6	1.0 ± 1.6	12.1 ± 2.6	2.0 ± 2.0	14.8 ± 2.7	23.6 ± 4.7	26.8 ± 9.3	25.1 ± 5.3
150	99.4 ± 0.4	98.2 ± 0.5	0.8 ± 1.1	11.7 ± 2.7	1.3 ± 1.3	13.5 ± 2.4	25.7 ± 5.5	27.5 ± 8.8	25.9 ± 4.9
160	99.5 ± 0.6	98.4 ± 0.5	0.8 ± 1.1	10.6 ± 3.0	1.3 ± 1.6	12.1 ± 3.0	25.9 ± 5.9	28.0 ± 9.1	26.1 ± 5.0
170	99.6 ± 0.5	98.7 ± 0.4	0.7 ± 1.0	9.5 ± 2.6	1.1 ± 1.3	10.8 ± 2.6	27.0 ± 6.5	28.8 ± 9.1	26.7 ± 5.2
180	99.7 ± 0.3	99.0 ± 0.3	0.6 ± 1.1	8.5 ± 3.1	0.9 ± 1.3	9.4 ± 3.1	27.8 ± 6.2	30.0 ± 8.6	28.0 ± 4.6
190	99.8 ± 0.3	99.2 ± 0.3	0.4 ± 0.7	7.6 ± 3.1	0.7 ± 0.9	8.4 ± 3.2	29.1 ± 6.2	30.9 ± 8.5	28.7 ± 4.4
200	99.8 ± 0.2	99.4 ± 0.2	0.2 ± 0.4	6.2 ± 2.8	0.4 ± 0.5	6.9 ± 2.8	30.3 ± 6.5	32.5 ± 8.3	31.0 ± 4.7

[a] Planning target volume (PTV) optimization priority

[b] Boost and elective PTVs

[c] Boost and elective PTV dose inhomogeneity

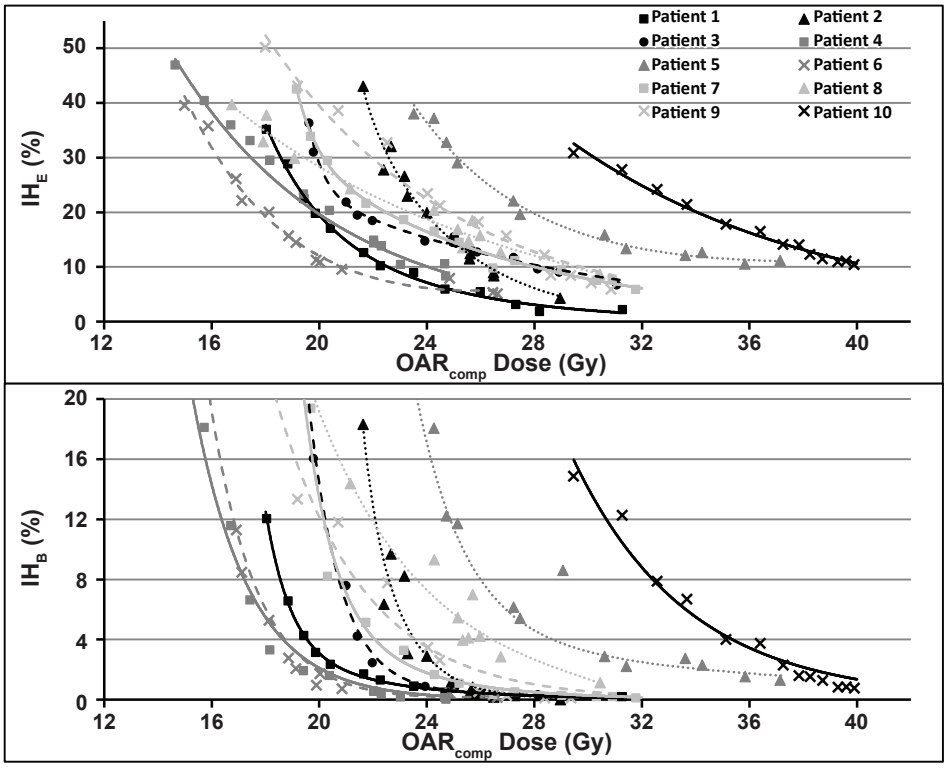


Figure 1. Plots of elective and boost planning target volume (PTV) inhomogeneity (IH_E [upper] and IH_B [lower]) and average mean composite organ at risk (OAR_{comp}) doses for multiple plans per patient in which the PTV optimization priority is systematically increased. Exponential fits are shown.

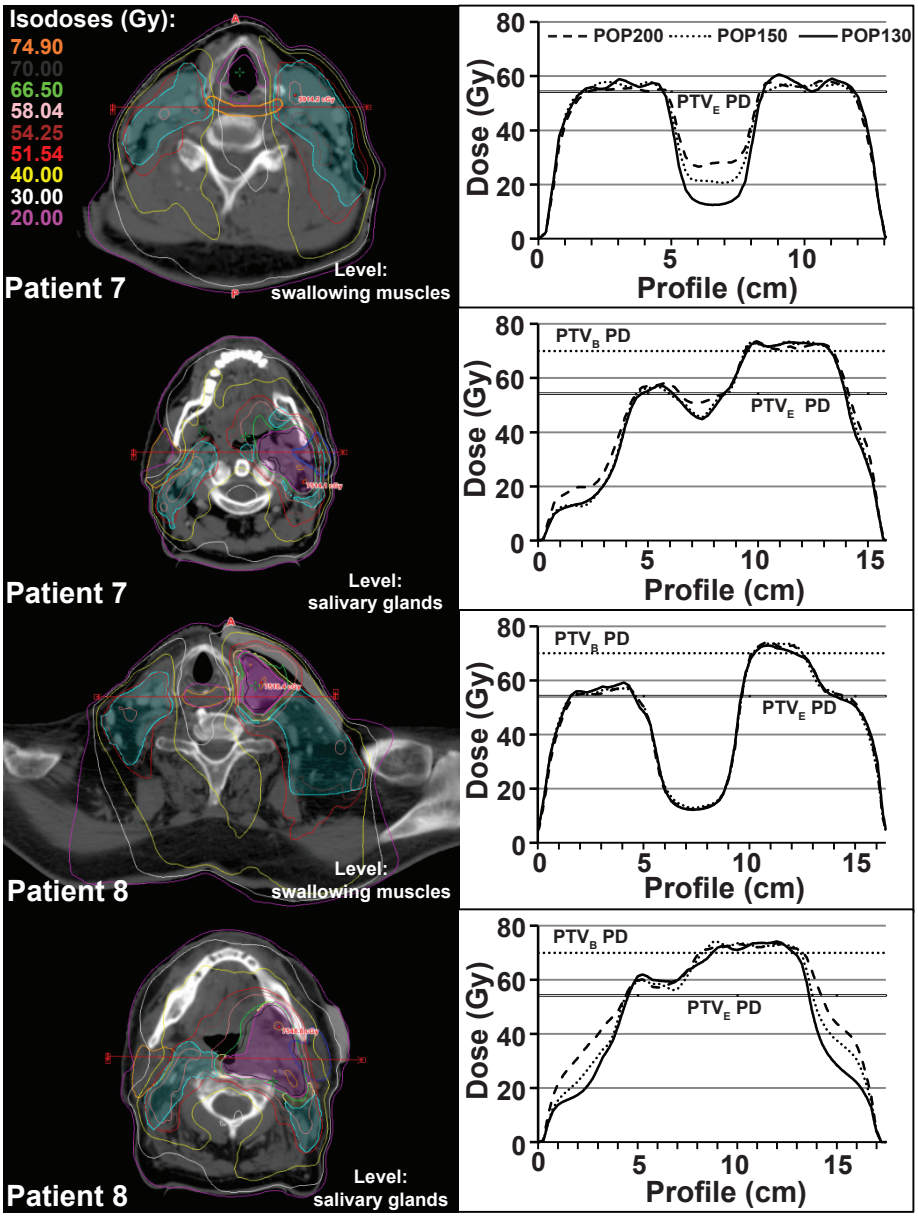


Figure 2. Dose distributions and profiles at the level of parotids (orange and blue contours) and cricopharyngeal muscle (orange contour) for patient 7 and 8. The red line indicates the location of the dose profile, shown by the graphs on the right. Elective and boost planning target volumes (PTV_B and PTV_E) are shown in cyan and magenta respectively, while their prescribed doses (PD) are indicated by the horizontal lines in the graphs.

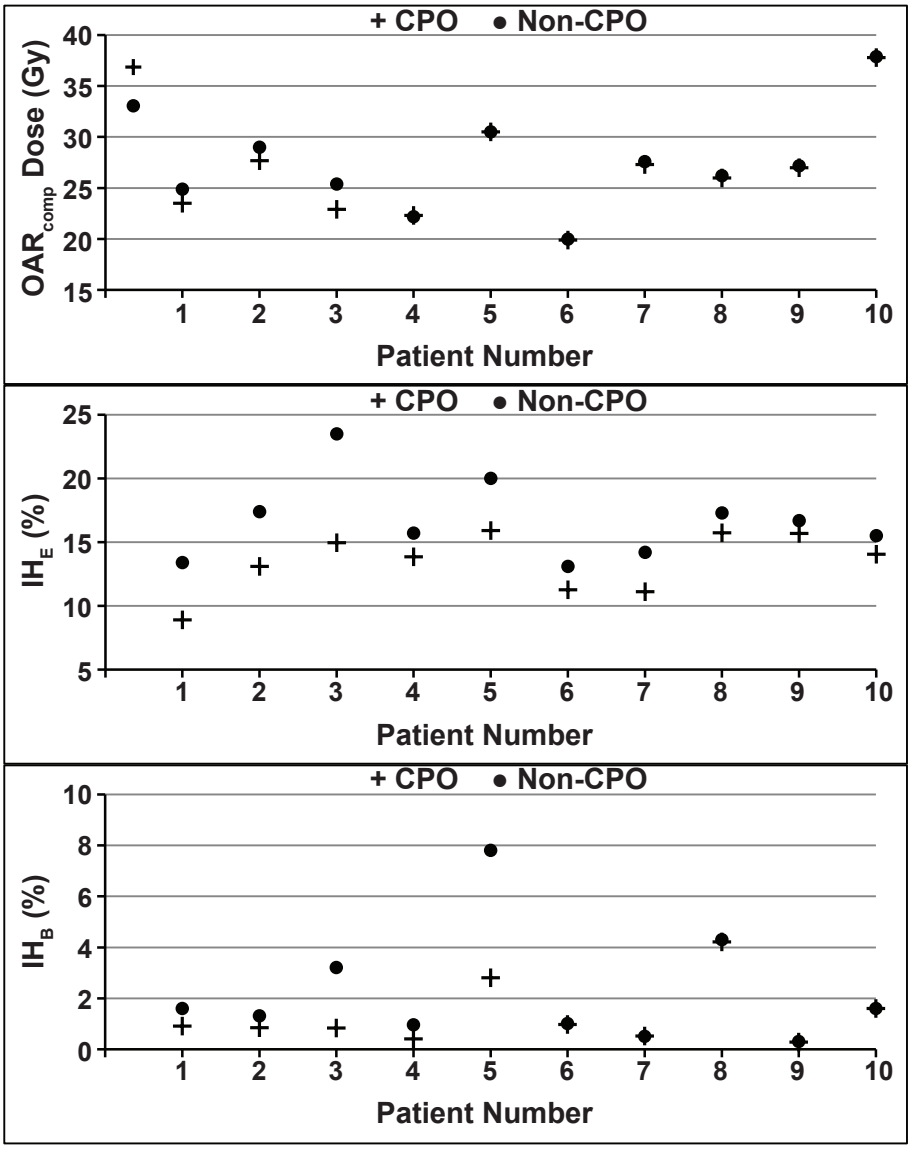


Figure 3. The results of continuous previous optimization (CPO) calculations on boost planning target volume (PTV) inhomogeneity (IH_B, lower), elective PTV inhomogeneity (IH_E, middle) and composite organ at risk dose (OAR_{comp}, average of the mean dose of the composite salivary and the mean dose of the composite swallowing structures) for every patient at PTV optimization priority of 150 (POP_{150}).

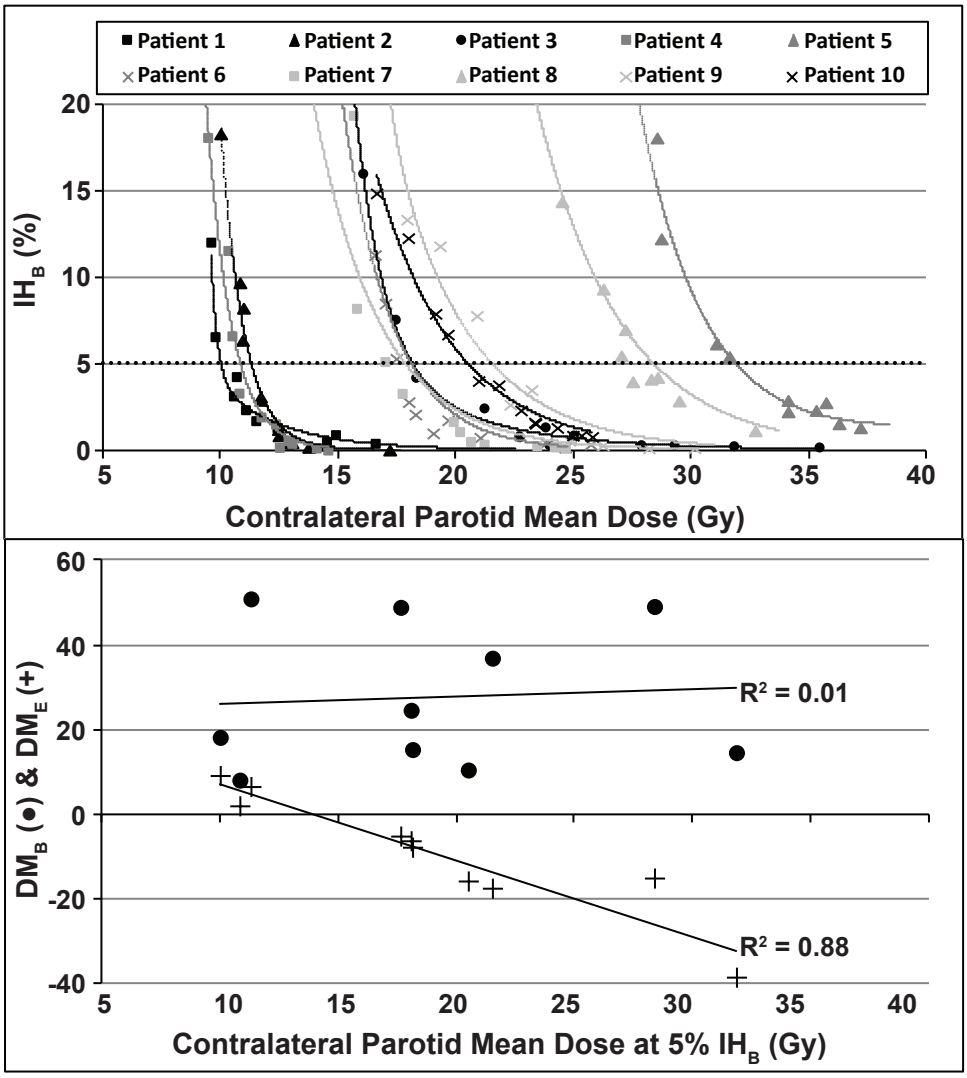


Figure 4. The upper graph shows the trade-off between the boost planning target volume (PTV) dose inhomogeneity (IH_B) and contralateral parotid mean dose for each patient with multiple plans per patient in which the PTV optimization priority is systematically increased. Exponential fits are shown. The intersections of the curves in the upper graph with the dotted line are used to obtain contralateral parotid mean dose at 5% IH_B. This is plotted against DM_E and DM_B in the lower graph for each patient. Linear fits are shown.

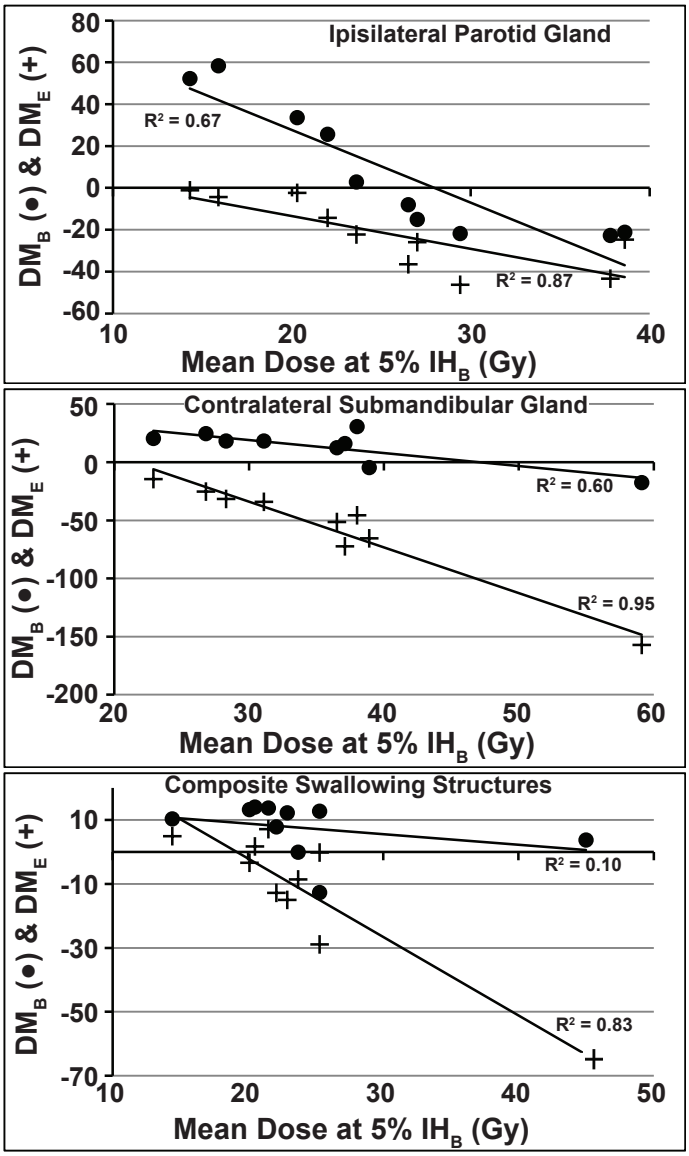


Figure 5. Mean dose to the ipsilateral parotid, contralateral submandibular gland and composite swallowing structures (x-axis) evaluated at 5% IH_B plotted against DM_B (circles, y-axis) and DM_E (plusses, y-axis) for all 10 patients plus their respective linear fits. The dose data points are obtained from similar trade-off curves as in the upper graph of Figure 4.

Discussion

This study quantitatively assessed the relationship between PTV dose homogeneity and OAR sparing in complex head and neck cancer VMAT plans. Excellent exponential fits were obtained in all patients between the average of the mean doses to the composite salivary and swallowing structures and dose inhomogeneity in the boost and elective PTVs (R^2 = 0.98±0.01 / 0.97±0.03, respectively). This was investigated for a range of patients with considerable variation in PTV size and location. However, the magnitude of OAR dose sparing for a given level of clinically relevant PTV inhomogeneity varied substantially between patients, showing a clear linear correlation with a modified Euclidean distance measure that was used to quantify the OAR-PTV_B / PTV_E distance. The results show that demanding highly homogeneous PTV dose coverage can considerably increase OAR doses, whereas reducing PTV dose homogeneity by, on average, fractions of a percent in the range of PTV_B V95% 99-100% enables substantial gains (in the order of several Gy) in average OAR sparing (Table 2). These are therefore important variables in developing automated planning solutions that can serve a variety of institutional plan acceptance criteria.

Similar exponential trade-offs are reported in investigations regarding Pareto-optimal planning[22–24]. These studies commonly assess the trade-off in the sparing of two OARs by changing the relative optimization weights of the OAR objectives using intensity modulated radiotherapy (IMRT). Most Pareto-optimal planning studies are not performed for head and neck cancer, and the ones that are use less complex planning conditions, such as a smaller number of OARs. Similar trade-offs, but between maximum PTV doses and OAR sparing, were reported by Craft et al.[25], although their data was derived using only two patients, and under more straightforward planning conditions. In contrast, we have used a bigger patient group of complex head and neck VMAT cases containing many individually constrained OAR (e.g. eight separate swallowing muscles and 3-4 salivary glands). Typical studies on the Pareto-front require the creation of >200 plans[9–11] of which the vast majority are not Pareto-optimal and therefore clinically less relevant. The excellent correlation coefficients for our exponential fits in each patient, suggests that we obtain near Pareto-optimal plan quality using our standard optimization procedure and a commercial planning system. Nonetheless, some data-points did deviate from the fit and this is likely due to differences in manual interactive optimization. The consistency of our institutional interactive optimization method, in spite of some differences due to the manual character, is also reflected by the good correlation of OAR doses with the modified Euclidean distance measure, based on only 10 data points per organ.

The clinical relevance of this work is further supported by differences in OAR sparing between clinical trial protocols. For example, amongst other objectives, a current RTOG head and neck study for oropharyngeal patients (*http://clinicaltrials.gov/show/NCT01302834*) states that 95% of the high-dose PTV should receive 100% of the prescribed dose, no more than 1cc may receive more than 110% of the prescription dose and no more than 0.03cc may receive less than 95% prescribed dose. In contrast, an upcoming EORTC (*http://clinicaltrials.gov/ct2/show/NCT01880359*) study states that the median PTV dose (D50%) should be the prescription dose (±2%), D95% / D98% of PTVs should receive ≥95% / ≥90% of the prescription dose and D5% of PTV_B should be ≤107% of prescription dose. The EORTC protocol is therefore generally less demanding than the RTOG requirements and substantial differences in OAR sparing (and differences in mean PTV dose) may be expected between these protocols. Indeed, after renormalization of the plan with the lowest POP value that satisfied the respective protocol and averaged over all patients, RTOG and EORTC acceptable plans had mean PTV doses of 73.0Gy and 71.0Gy, respectively, and mean doses to the composite salivary and swallowing structures were on average 8.0Gy and 5.9Gy lower in EORTC plans compared to RTOG plans. These differences impose limitations on the comparison of treatment plans and clinical results created using different planning protocols, and should be taken into consideration in the design of knowledge-based planning strategies. One example would be that if libraries consisting of previous clinical plans are used to generate plans for new patients[6,13,26–28] then it is likely that such libraries will contain plans with PTV dose homogeneity preferred by an individual center. This may limit library sharing amongst institutions that have a preference for different levels of PTV dose homogeneity.

In comparison with the available literature, our study is distinguished by the number of OAR, treatment plans and patients. In a limited study using IMRT, de Kruijf et al.[29] reported trade-offs between parotid gland sparing and PTV underdosage in 4 oropharynx cancer patients (16 plans). They found that their PTV underdosage could facilitate a decrease in mean dose to parotid glands of >10Gy. Houweling et al.[30] showed in 10 oropharynx cancer patients (40 plans) that by reducing the coverage of 99% of the contralateral elective PTV from 95 to 90% of the prescribed dose, the mean dose to the contralateral submandibular gland could be reduced from 54Gy to approximately 40Gy. However, these PTV dose homogeneity values are unlikely to be accepted in clinical practice. In contrast, the present data emphasizes the potential for improvements in OAR sparing within the typical range of clinically acceptable levels of PTV dose coverage / homogeneity.

Finally, the present study includes a further dimension in that we have also examined the relationship between the PTV-OAR dose trade-off and the PTV–OAR distance (DM_B / DM_E). This has shown that distance to PTV_E primarily contributes to the dose of the majority of the OAR, whereas distance to PTV_B was of limited influence, suggesting that increased sparing of these OAR may therefore be achieved by solely reducing PTV_E dose homogeneity whilst preserving PTV_B coverage. Such strategies, including the steering of inhomogeneity into specific areas of the elective PTV should also take into account clinical patterns of failure. The relation between DM_B / DM_E and OAR_{comp} dose was investigated but did not correlate well for the entire composite OAR because different OAR typically have differently shaped dose distributions due to variation in the distances between them and PTV_B / PTV_E. This is demonstrated by the different slopes of the linear fits in Figure 4. An investigation using more patient data with varying geometry should demonstrate whether the relationship between DM_B / DM_E and mean OAR dose for a large range of IH_B / IH_E values can be used to estimate achievable OAR sparing for a chosen level of PTV dose homogeneity and avoid the need to generate multiple plans. Several studies using static-field IMRT[20,31] report on the relation between PTV-OAR distance and achievable OAR dose. Using IMRT, however, the distance measure of out-of-field voxels should be increased to compensate for their lower dose. Such voxels do not exist in the axial planes containing PTV using VMAT because of the full arc irradiation, thereby hampering comparison with results obtained using static IMRT[32].

Potential limitations of this study include the following. Differences in commercial TPS and VMAT systems may affect extrapolation of these results. However, testing multiple planning systems was beyond the scope of this particular study. We were not seeking to describe the inner workings of the TPS, which was simply used as a tool with which to investigate PTV-OAR dose trade-offs. Excluding OARs with large PTV overlap gave an inconsistent set of OARs amongst patients, which influenced the obtained distance measures. In addition, the adjustment factor for dose in voxels cranio-caudal to the total PTV was chosen arbitrarily.

In conclusion, within clinically relevant ranges, OAR sparing is highly dependent on the permitted boost and, to a higher extent, elective PTV dose inhomogeneity and distance. These results are relevant for future automated treatment planning strategies, clinical practice, and when comparing plans and outcomes according to different clinical planning protocols.

References

1 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

2 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

3 E.M. Quan, J.Y. Chang, Z. Liao, T. Xia, Z. Yuan, H. Liu, X. Li, C.A. Wages, R. Mohan, and X. Zhang, "Automated volumetric modulated Arc therapy treatment planning for stage III lung cancer: how does it compare with intensity-modulated radio therapy?," Int. J. Radiat. Oncol. Biol. Phys. **84**(1), e69–76 (2012).

4 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **79**(4), 1241–7 (2011).

5 B. Wu, D. Pang, P. Simari, R. Taylor, G. Sanguineti, and T. McNutt, "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study.," Med. Phys. **40**(2), 021714 (2013).

6 M. Kazhdan, P. Simari, T. McNutt, B. Wu, R. Jacques, M. Chuang, and R. Taylor, "A shape relationship descriptor for radiation therapy planning.," Med. Image Comput. Comput. Assist. Interv. **12**(Pt 2), 100–8 (2009).

7 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

8 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

9 T. Janssen, Z. van Kesteren, G. Franssen, E. Damen, and C. van Vliet-Vroegindeweij, "Pareto fronts in clinical practice for pinnacle.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 873–80 (2013).

10 Z. van Kesteren, T.M. Janssen, E. Damen, and C. van Vliet-Vroegindeweij, "The dosimetric impact of leaf interdigitation and leaf width on VMAT treatment planning in Pinnacle: comparing Pareto fronts.," Phys. Med. Biol. **57**(10), 2943–52 (2012).

11 A.L. Hoffmann, A.Y.D. Siem, D. den Hertog, J.H.A.M. Kaanders, and H. Huizenga, "Derivative-free generation and interpolation of convex Pareto optimal IMRT plans.," Phys. Med. Biol. **51**(24), 6349–69 (2006).

12 M. Steneker, A. Lomax, and U. Schneider, "Intensity modulated photon and proton therapy for the treatment of head and neck tumors.," Radiother. Oncol. **80**(2), 263–267 (2006).

13 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

14 K.L. Moore, R.S. Brame, D.A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **81**(2), 545–51 (2011).

15 C. Thilmann, A. Zabel, S. Nill, B. Rhein, A. Hoess, P. Haering, S. Milke-Zabel, W. Harms, W. Schlegel, M. Wannenmacher, and J. Debus, "Intensity-modulated radiotherapy of the female breast.," Med. Dosim. **27**(2), 79–90 (2002).

16 M.E.M.C. Christianen, J.A. Langendijk, H.E. Westerlaan, T.A. van de Water, and H.P. Bijl, "Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.," Radiother. Oncol. **101**(3), 394–402 (2011).

17 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–9 (2009).

18 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

19 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

20 M. Kazhdan, T. McNutt, R. Taylor, B. Wu, and P. Simari, "Comment on 'A planning quality evaluation tool for prostate adaptive IMRT based on machine learning' [Med. Phys. **38**, 719 (2011)].," Med. Phys. **38**(5), 2820; author reply 2821 (2011).

21 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

22 D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld, "Multicriteria VMAT optimization.," Med. Phys. **39**(2), 686–96 (2012).

23 M. Monz, K.H. Küfer, T.R. Bortfeld, and C. Thieke, "Pareto navigation: algorithmic foundation of interactive multi-criteria IMRT planning.," Phys. Med. Biol. **53**(4), 985–98 (2008).

24 R.O. Ottosson, P.E. Engstrom, D. Sjöström, C.F. Behrens, A. Karlsson, T. Knöös, and C. Ceberg, "The feasibility of using Pareto fronts for comparison of treatment planning systems and delivery techniques.," Acta Oncol. **48**(2), 233–7 (2009).

25 D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning.," Med. Phys. **33**(9), 3399–407 (2006).

26 D. Good, J. Lo, W.R. Lee, Q.J. Wu, F.-F. Yin, and S.K. Das, "A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning.," Int. J. Radiat. Oncol. Biol. Phys. **87**(1), 176–81 (2013).

27 B. Wu, T. McNutt, M. Zahurak, P. Simari, D. Pang, R. Taylor, and G. Sanguineti, "Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **84**(5), e647–53 (2012).

28 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

29 W.J.M. de Kruijf, B.J.M. Heijmen, and P.C. Levendag, "Quantification of trade-off between parotid gland sparing and planning target volume underdosages in clinically node-negative head-and-neck intensity-modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **68**(1), 136–43 (2007).

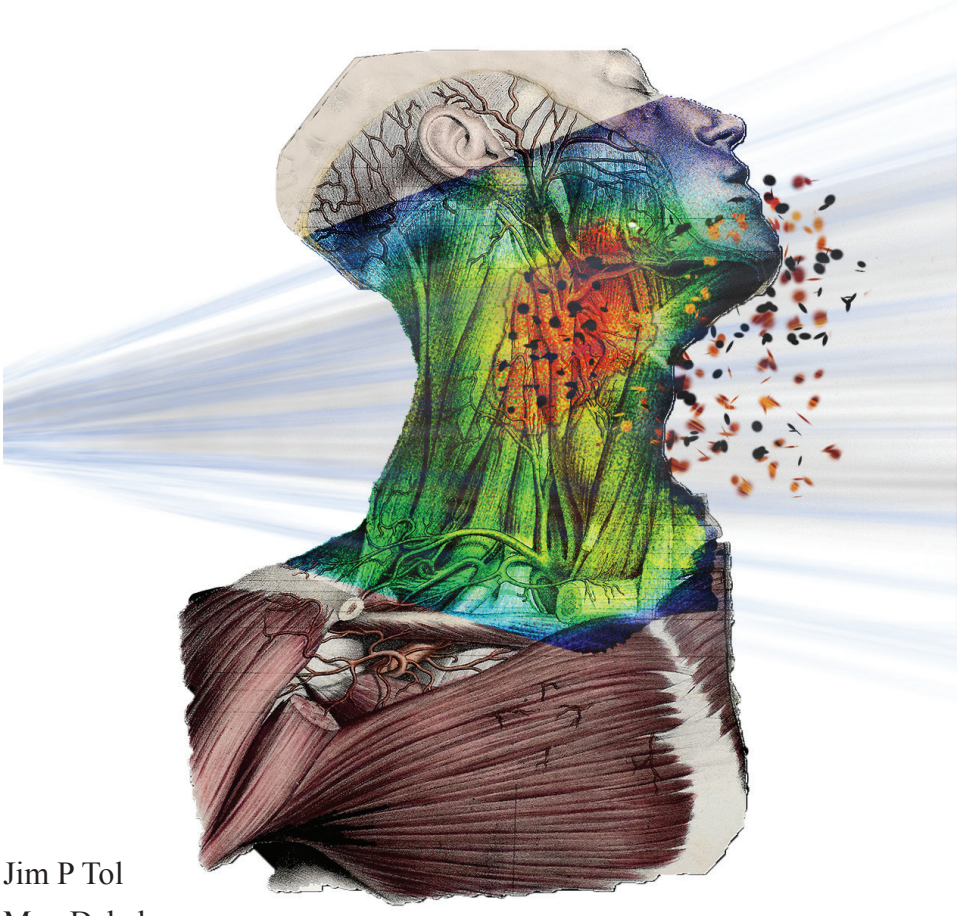
30 A.C. Houweling, T. Dijkema, J.M. Roesink, C.H.J. Terhaard, and C.P.J. Raaijmakers, "Sparing the contralateral submandibular gland in oropharyngeal cancer patients: a planning study.," Radiother. Oncol. **89**(1), 64–70 (2008).

31 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt, "Patient geometry-driven information retrieval for IMRT treatment plan quality control.," Med. Phys. **36**(12), 5497–505 (2009).

32 S.F. Petit, B. Wu, M. Kazhdan, A. Dekker, P. Simari, R. Kumar, R. Taylor, J.M. Herman, and T. McNutt, "Increased organ sparing using shape-based treatment plan optimization for intensity modulated radiation therapy of pancreatic adenocarcinoma.," Radiother. Oncol. **102**(1), 38–44 (2012).

Chapter 4

Different treatment planning protocols can lead to large differences in organ at risk sparing



Jim P Tol
Max Dahele
Patricia Doornaert
Ben J Slotman
Wilko FAR Verbakel

Radiotherapy and Oncology **113**(2), 267-271.

Abstract

Purpose

Different treatment planning protocols may define varying planning target volume (PTV) dose criteria. We investigated the hypothesis that this could result in differences in organ-at-risk (OAR) sparing.

Materials and Methods

Volumetric modulated arc therapy (VMAT) plans were created for ten locally advanced head and neck cancer patients following PTV dose criteria specified by the RTOG, EORTC and institutional (VUMC) protocols. Resulting plans were evaluated on the basis of the homogeneity index, calculated for the boost / elective PTVs as HI_B / HI_E =100% × (D2% – D98%)/D50% and mean dose to the individual and composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) OARs.

Results

RTOG plans were the most homogeneous, with a mean HI_B of 8.2±0.9%, compared to 9.5±1.0% / 11.6±1.5% for the VUMC / EORTC plans. EORTC plans provided the most OAR sparing, with $comp_{sal}$ / $comp_{swal}$ doses of 24.6±7.7Gy / 22.9±4.2Gy, compared to 32.2±9.7Gy / 29.9±4.2Gy and 28.4±8.1Gy / 24.7±5.3Gy for RTOG and VUMC, respectively. EORTC provided 7.2Gy / 7.7Gy mean dose reductions to the contralateral / ipsilateral parotid glands compared to RTOG.

Conclusions

Different planning protocols resulted in different levels of PTV dose homogeneity. We observed differences ≥7Gy in composite and individual mean OAR doses. This could influence rates of toxicity and should be taken into account when comparing clinical studies. A consensus should therefore be reached between major trial groups on appropriate PTV parameters.

Introduction

Intensity modulated radiotherapy (IMRT) and its rotational variant volumetric modulated arc therapy (VMAT), are increasingly used to improve organ-at-risk (OAR) sparing while maintaining adequate and homogeneous planning target volume (PTV) coverage. Treatment planning for IMRT and VMAT typically involves trying to meet specified criteria for OAR sparing and PTV dose coverage and homogeneity. It may involve the prioritization of certain OARs above others or balancing OAR sparing against PTV dose homogeneity[1–4].

Different institutional and international study protocols may define varying OAR and PTV dose criteria for the same type of treatment. This also holds for two major trials for locally advanced head and neck cancer (HNC); one from the Radiation Therapy Oncology Group (RTOG, Clinicaltrials.gov No. NCT01302834), and another from the European Organization for Research and Treatment of Cancer (EORTC, Clinicaltrials.gov No. NTC01880359). Both protocols also differ from our current institutional (VUmc) HNC planning protocol. Based on previous work[4], such variation in planning criteria for PTV dose homogeneity was expected to result in differences in OAR sparing between the planning protocols and could influence expected rates of toxicities such as xerostomia and dysphagia. This would be relevant to interpreting and comparing the results of clinical studies. To the best of our knowledge, the extent of differences in OAR sparing following commonly used clinical planning protocols has not been quantified in the literature. To investigate this we compared OAR sparing and PTV dose homogeneity when planning HNC patients with the PTV dose criteria specified by RTOG, EORTC and VUmc protocols.

Materials and Methods

General Considerations

Planning CT scans with a 2.5mm slice thickness from 10 locally advanced HNC patients previously treated using RapidArc™ (Varian Medical Systems, Palo Alto, USA) were arbitrarily selected for this retrospective planning study. Contouring of the tumor and affected lymph nodes was done using all available diagnostic information; MRI and PET-CT scans were (rigidly) co-registered. Elective lymph node regions were contoured following the consensus guidelines[5,6] and the swallowing muscles were delineated according to Christianen et al.[7].

Plans were created using our current institutional simultaneous integrated boost (SIB) technique which has been described in more detail previously[4,8,9]. In brief, this aims to deliver 54.25Gy / 70Gy to the elective / boost PTV (PTV_E / PTV_B) in 35 fractions of 1.55Gy and

2Gy, respectively. The clinical target volume (CTV) was defined as a 5mm expansion of the gross tumor volume and biopsy proven positive lymph node regions, edited for anatomical boundaries. PTV_B was obtained by a 4-5mm expansion of the CTV. PTV_E consisted of a 4-5mm expansion of the elective nodal regions, minus PTV_B and minus a 5mm transition zone (PTV_T) from PTV_B to PTV_E, which was created to allow for a dose fall-off between these PTVs. For optimization and reporting purposes only, a 'virtual' build-up region was routinely created where the PTV extended outside the body by expanding the body 5mm beyond the PTV. All dosimetric parameters were reported using this structure. Table 1 shows the size of the PTVs, along with tumor site and stage for all patients.

During optimization, maximum dose objectives were used for the spinal cord, brainstem and their planning at risk (PRV) volumes after a 3mm expansion. Since shoulder position can vary slightly during and between fractions, a maximum dose objective was used to limit direct irradiation through them. For salivary and swallowing OARs, 3 to 5 optimization objectives, placed evenly along the dose-volume histogram (DVH)-line, were used to reduce their mean dose. Throughout the RapidArc optimization process, the planner interactively adapted the objectives with the aim of keeping them at a fixed distance to the DVH-line[4,8,9]. Optimization was typically performed using the part of the OAR outside the PTV.

Table 1. The disease site, stage and volumes of the elective, transition and boost planning target volumes (PTV_E, PTV_T, PTV_B, respectively) and the volumes of the composite salivary ($comp_{sal}$) and swallowing structures ($comp_{swal}$) for all patients.

Patient Number	Disease Site	Stage	PTV_E (cm^3)	PTV_T (cm^3)	PTV_B (cm^3)	$comp_{sal}$ (cm^3)	$comp_{swal}$ (cm^3)
1	Oropharynx	T2N0	346.6	70.2	28.3	41.7	17.3
2	Oropharynx	TxN2a	240.6	43.4	94.5	42.4	25.0
3	Nasopharynx	T2N0	538.1	33.9	123.5	105.4	31.7
4	Supraglottic larynx	T3N2b	358.5	80.1	143.4	42.4	9.9
5	Oropharynx	T3N2c	433.8	103.0	313.0	74.0	17.3
6	Oropharynx	T2N2a	592.6	104.6	175.6	90.1	12.4
7	Oropharynx	T2N2b	441.7	48.3	188.6	82.3	16.5
8	Oropharynx	T2N2b	457.9	34.4	289.0	70.4	18.5
9	Oropharynx	T4bN0	422.7	58.6	117.7	86.8	49.6
10	Supraglottic larynx	T2N2c	428.7	93.7	243.3	66.7	7.1

Depending on the degree of overlap with the PTVs and choice for inclusion made by the treating clinician, the OARs could include some or all of the salivary glands (contralateral and ipsilateral parotid and submandibular glands, with laterality determined by the dominant location of the boost PTV) and swallowing muscles[7] (upper esophageal sphincter, supraglottic and glottic larynx, superior, medial and inferior pharyngeal constrictor muscle and the cricopharyngeal muscle). Composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) structures were created to simplify the dose reporting. OAR_{comp} mean dose was obtained by averaging the mean doses to $comp_{sal}$ and $comp_{swal}$. Table 1 shows the size of $comp_{sal}$ and $comp_{swal}$ for all patients.

Planning Protocols

For each patient, three different treatment plans for the previously described dose prescription were created following the RTOG, EORTC, and our institutional protocol. Optimization was performed using the progressive resolution optimizer (PRO) v10.0.28 in the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, USA) v10. The normal tissue objective (NTO) function with same settings for all plans was routinely used during optimization. Dose calculation, with a subsequent 'continue previous optimization' (CPO) to improve PTV dose homogeneity[4], was performed with a 2.5mm grid resolution using the anisotropic analytical algorithm (AAA) v10.0.28, including tissue heterogeneity corrections. The PTV optimization objective weightings were varied in order to meet the criteria for PTV dose coverage and homogeneity specified by the respective protocol. Optimization weightings for the OAR objectives were kept constant between the different protocol plans. Using the approach of interactive optimization described above, the resulting plans show consistent exponential trade-offs between OAR sparing and PTV dose homogeneity[4]. To obtain optimal OAR sparing for each protocol, plans were created that only just satisfied the specified dose criteria for the PTVs.

(i) ***RTOG:*** In an ongoing RTOG HNC study for oropharyngeal patients (*http://clinicaltrials.gov/show/NCT01302834*), 95% of PTV_B should receive 100% of the prescribed dose (PD, D95% = 100%), no more than $1cm^3$ may receive more than 110% of the PD (V110% < $1cm^3$) and no 'cold spots' <95% PD may be present greater than $0.03cm^3$ in size. (ii) ***EORTC:*** In a soon to be recruiting EORTC study for HNC patients (*http://clinicaltrials.gov/ct2/show/NCT01880359*), D95% / D98% of the PTVs should receive >95% / >90% of the PD, D5% of PTV_B should be ≤107% of PD, and the median PTV dose (D50%) should be the PD±2%. (iii) ***VUMC:*** Our current institutional planning protocol for HNC aims for PTV_B / PTV_E to receive V95% > 99% / 98% and V107% < 4% / 15%, although

exceptions can be made for individual patients. Hot spots, defined as regions with >107% of the PD, are to be avoided in the mandible or laryngeal cartilage. The resulting mean PTV_B dose should typically be between 101-103% of the prescribed dose.

Study Endpoints

The different protocol plans were compared on the basis of (i) mean dose to individual and composite salivary and swallowing OARs and a composite of all OARs, and (ii) the homogeneity index (HI) for PTV_B / PTV_E, calculated using HI_B / HI_E = 100% × (D2% – D98%) / D50%. Statistically significant differences ($p \leq 0.05$) between the plans made following the RTOG, EORTC and VUMC planning protocols were determined using paired two-sided Student t-tests.

Results

Clinically acceptable maximum doses to the brainstem and spinal cord were achieved in all plans. Results obtained when planning following the three protocols were averaged over all 10 patients and are summarized in Table 2. Different levels of OAR sparing and PTV dose homogeneity were obtained in the three protocols. All plans satisfied the respective protocols regarding PTV dose coverage and homogeneity. RTOG plans were the most homogeneous, obtaining HI_B / HI_E values of 8.2±0.9% / 12.3±1.2%, compared to 9.5±1.0% / 15.1±0.7% and 11.6±1.5% / 19.6±1.3% for the VUMC and EORTC plans, respectively. PTV_B dose coverage also decreased from the RTOG to VUMC and EORTC plans, with PTV_B V95% values of 100.0±0.0%, 99.4±0.5% and 97.8±1.2% obtained, respectively. All plans had a boost CTV coverage of 100% of the PD, with the exception of one EORTC plan (CTV V95% = 99.7%)

Consistent with more homogeneous target dose coverage, the RTOG plans had the highest OAR doses, with a $comp_{sal}$ mean dose of 32.2±9.7Gy on average, 3.8Gy / 7.6Gy higher than in the VUMC / EORTC plans. Differences in $comp_{swal}$ mean dose were even higher, with average values of 29.9±4.2, 24.7±5.3 and 22.9±4.2Gy obtained following the RTOG, VUMC and EORTC protocols, respectively. Mean dose differences between the protocols were similar for the individual salivary glands (Table 2).

The penalty for obtaining a more homogeneous PTV dose coverage on OAR sparing varied between patients, with the range of mean dose differences being 1.5-17.3Gy and 2.2-15.7Gy for the contralateral and ipsilateral parotid glands, respectively, between the RTOG and EORTC protocol plans. The trade-off between boost / elective PTV dose homogeneity (HI_B / HI_E) and OAR_{comp} sparing in the different protocol plans is shown for each patient in Figure 1. The RTOG plans (red) group together at lowest HI values and consistently

result in higher OAR_{comp} doses than the EORTC (blue) or VUmc (green) plans. Figure 2 shows a typical example of the dose range where OAR DVHs improve by allowing more inhomogeneous PTV doses.

Because the RTOG plans are normalized to deliver a V100% of 95%, their mean PTV_B doses were slightly higher (72.6±0.4%), compared to the VUMC / EORTC plans (71.7±0.3% / 71.4±0.2%). In Table 2, all differences between planning protocols (RTOG-EORTC, RTOG-VUMC and EORTC-VUMC) were significant, except for the difference in mean dose to the superior pharyngeal constrictor muscle (PCM) and $comp_{swal}$ between the EORTC-VUMC protocol plans, and the mean dose to the supraglottic larynx in the RTOG-VUMC plans.

Table 2. Planning target volume (PTV) dose coverage and homogeneity values and mean doses to individual and composite salivary and swallowing organs-at-risk (OARs) when planning according to three clinical protocols, averaged over all ten patients.

Planning Protocol	RTOG	VUMC	EORTC
Boost PTV			
V95%[a] (%)	100 ± 0.0	99.4 ± 0.5	97.8 ± 1.2
Mean Dose (Gy)	72.6 ± 0.4	71.7 ± 0.3	71.4 ± 0.2
HI_B[b] (%)	8.2 ± 0.9	9.5 ± 1.0	11.6 ± 1.5
HI_E[b] (%)	12.3 ± 1.2	15.1 ± 0.7	19.6 ± 1.3
Mean dose (Gy)			
CL Parotid	25.6 ± 8.7	20.9 ± 6.6	18.4 ± 6.1
IL Parotid	31.9 ± 8.2	27.0 ± 6.1	24.2 ± 6.2
CL SMG[c]	38.8 ± 7.8	34.5 ± 5.7	31.4 ± 5.0
$comp_{sal}$[d]	32.2 ± 9.7	28.4 ± 8.1	24.6 ± 7.7
Supraglottic Larynx	34.1 ± 5.5	31.9 ± 6.3	28.7 ± 5.1
Superior PCM[e]	35.4 ± 19.4	33.4 ± 18.8	29.6 ± 18.7
$comp_{swal}$[d]	29.9 ± 4.2	24.7 ± 5.3	22.9 ± 4.2
OAR_{comp}[f] (Gy)	31.1 ± 4.8	26.6 ± 4.6	23.8 ± 4.0

[a] Volume of boost PTV receiving ≥95% prescribed dose

[b] Boost and elective PTV homogeneity indices

[c] Contralateral submandibular gland

[d] Composite salivary and swallowing structures

[e] Pharyngeal constrictor muscle

[f] Average of the mean doses to the composite salivary and swallowing structures

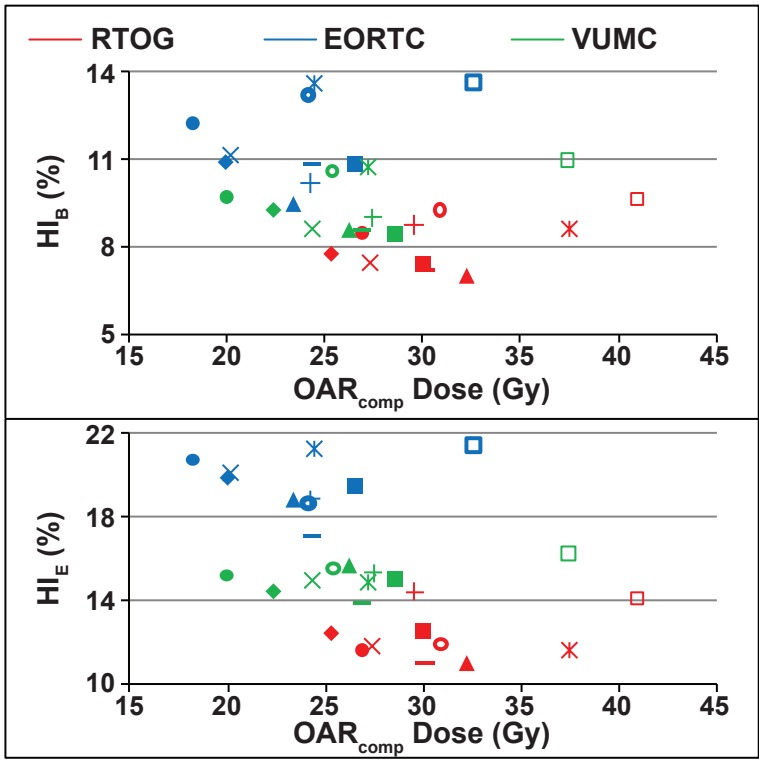


Figure 1. The trade-off between PTV dose homogeneity (HI_B / HI_E) and composite OAR (OAR_{comp}) mean dose for the different patients (symbols) planned according to the EORTC (blue), VUMC (green) and RTOG (red) protocols.

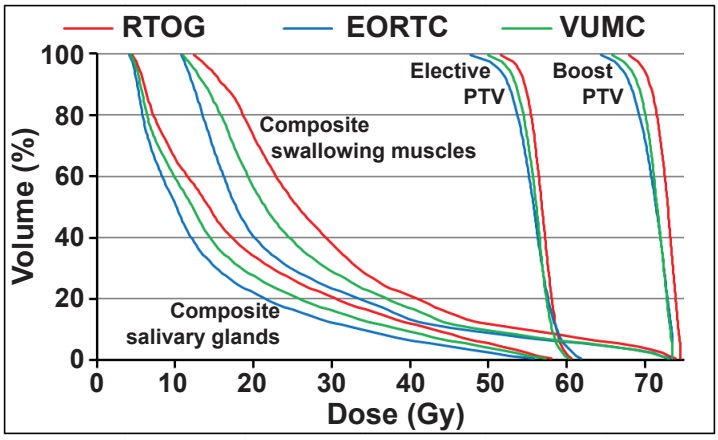


Figure 2. Dose-volume histogram results for a typical patient planned according to the EORTC (blue), VUMC (green) and RTOG (red) protocols.

Discussion

This study showed that OAR sparing can differ by more than 7Gy on average depending on the PTV dose homogeneity criteria that are used to create HNC plans according to three different clinical planning protocols. The most homogeneous PTV doses were achieved using the RTOG protocol whereas the EORTC PTV criteria led to the lowest PTV dose homogeneity. In line with previous research[4], these relatively modest differences had a large impact on OAR sparing, with for example contralateral and ipsilateral parotid gland mean doses of 7.2 and 7.7Gy lower on average for EORTC plans compared to RTOG plans. As expected based on the PTV planning criteria, results obtained when using our institutional protocol fell between those for the RTOG and EORTC. To the best of our knowledge, this variation in planning criteria following the application of different study and institutional protocols and the subsequent effect of this on OAR dose is not an issue that has previously been highlighted or adequately addressed in the existing literature.

Regarding recent literature on normal tissue complication probability (NTCP) modeling of xerostomia[10–13] and dysphagia[14–16], the observed differences in OAR sparing are clinically relevant and indicate that the chosen planning protocol may influence expected salivary and swallowing toxicity. Dijkema et al. analyzed the combined Michigan and Utrecht data and found a stretched out sigmoidal NTCP curve correlating with mean parotid dose[13]. Judging from their data, reducing parotid gland mean doses by 7.2-7.7Gy, as seen in this study, may reduce NTCP (i.e. a reduction of the salivary flow ratio to <25%) values by up to approximately 20%, depending on the initial dose. Although we aimed to reduce dose to all swallowing muscles, Christianen et al.[16] found that mean doses to the supraglottic larynx and superior pharyngeal constrictor muscle were the most predictive of swallowing dysfunction after HNC radiotherapy treatment. In the present study, averaged mean doses to these structures ranged from 28.7-34.1Gy and 29.6-35.4Gy in the different protocols (Table 2). Depending on the initial dose, such improvements in OAR sparing following the EORTC protocol, could lead to relevant improvements in NTCP. Time and resources are being spent into furthering the development of novel, and expensive treatment techniques, with the aim of improving OAR sparing over conventional radiotherapy treatments. However, the present results show that large gains in OAR sparing using existing photon radiotherapy can also be obtained by treating patients according to a clinical protocol that allows for a slightly lower PTV dose coverage and homogeneity. Attempts to increase OAR sparing often result in dose distributions that tightly curve around the critical structures. Consistent and accurate contouring of the elective node regions and OARs is therefore important to achieve optimal OAR sparing and to facilitate accurate outcome reporting and comparison between studies.

In terms of whether the less stringent PTV criteria for the EORTC protocol might lead to a less effective treatment, we are not aware of any studies that had local control as an a priori endpoint and investigated its relationship to varying PTV coverage in the ranges described in this study. Nonetheless there is some circumstantial indication from the literature that the small differences in PTV coverage / dose observed in this study may not translate into easily detected differences in local control, including: (1) Several studies report on recurrence patterns in head and neck cancer following IMRT. Tumor recurrences occur in the vast majority of cases within the GTV; although different methods of reconstructing the received dose have their shortcomings, it is reasonable to assume that in such cases the area of recurrence has received (close to) the planned prescription dose[17–19] and radioresistance, rather than marginal dose differences, may be the most likely cause of treatment failure[17,18]; (2) In the PARSPORT study comparing 3D conformal radiotherapy (3D-CRT) and IMRT for HNC, no difference in local control was observed at 24 months[20]. This is despite a planning study (n=5 patients) from the trial group showing that minimum dose regions located in or near the match regions were improved by using IMRT instead of 3DCRT, along with tumor coverage, dose homogeneity and conformality. Although these differences were not statistically significant and the sample size was small, this led the authors to suggest that improved dosimetry with IMRT should translate into improved local control[21]. At the present time, there remains a lack of data showing that using RTOG-like PTV metrics results in clinically relevant increases in tumor control when compared with an EORTC-like approach. If this is indeed the case, using an EORTC-like protocol may realize substantial gains in OAR dosimetry without compromising local control.

One potential limitation of this study is that it was performed using HNC patients only. However, since the trade-off between PTV dose homogeneity and OAR sparing has also been shown in different tumor sites[1–3], in which differences in planning protocols also exist, the present observations are expected to hold regardless of tumor location. In addition, the magnitude of sparing using different protocols differed between patients. However, due to the limited number of included patients, we did not attempt to investigate whether these differences could correlate with HNC tumor site and PTV size. Although all plans were created in Eclipse we expect that the results are independent of the TPS. It should be noted that the CPO function of Eclipse to improve PTV dose homogeneity is also present in some other TPSs (e.g. named ‘warm start’ in the Pinnacle TPS [Philips Healthcare]).

In conclusion, because different planning protocols demand varying levels of PTV dose homogeneity, the chosen planning protocol can substantially influence the level of OAR sparing that can be achieved. This could influence rates of toxicity and should be taken into account when comparing clinical studies that have been performed using different planning protocols. The large differences in OAR sparing between different protocols suggest that a consensus on planning criteria for PTV dose coverage and homogeneity should be reached between organizations such as the RTOG and EORTC. This would ideally also include a consensus on contouring[5,6].

References

1 D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, “Approximating convex pareto surfaces in multiobjective radiotherapy planning.,” Med. Phys. **33**(9), 3399–407 (2006).

2 T. Janssen, Z. van Kesteren, G. Franssen, E. Damen, and C. van Vliet-Vroegindeweij, “Pareto fronts in clinical practice for pinnacle.,” Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 873–80 (2013).

3 Z. van Kesteren, T.M. Janssen, E. Damen, and C. van Vliet-Vroegindeweij, “The dosimetric impact of leaf interdigitation and leaf width on VMAT treatment planning in Pinnacle: comparing Pareto fronts.,” Phys. Med. Biol. **57**(10), 2943–52 (2012).

4 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, “Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity,” Med. Phys. **41**(2), 021722 (2014).

5 V. Grégoire, P. Levendag, K.K. Ang, J. Bernier, M. Braaksma, V. Budach, C. Chao, E. Coche, J.S. Cooper, G. Cosnard, A. Eisbruch, S. El-Sayed, B. Emami, C. Grau, M. Hamoir, N. Lee, P. Maingon, K. Muller, and H. Reychler, “CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC,RTOG consensus guidelines,” Radiother. Oncol. **69**(3), 227–236 (2003).

6 V. Grégoire, K. Ang, W. Budach, C. Grau, M. Hamoir, J.A. Langendijk, A. Lee, Q.-T. Le, P. Maingon, C. Nutting, B. O’Sullivan, S. V Porceddu, and B. Lengele, “Delineation of the neck node levels for head and neck tumors: a 2013 update. DAHANCA, EORTC, HKNPCSG, NCIC CTG, NCRI, RTOG, TROG consensus guidelines.,” Radiother. Oncol. **110**(1), 172–81 (2014).

7 M.E.M.C. Christianen, J.A. Langendijk, H.E. Westerlaan, T.A. van de Water, and H.P. Bijl, “Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.,” Radiother. Oncol. 101(3), 394–402 (2011).

8 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, “RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.,” Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

9 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, “Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.,” Radiat. Oncol. **6**(1), 74 (2011).

10 I. Beetz, C. Schilstra, A. van der Schaaf, E.R. van den Heuvel, P. Doornaert, P. van Luijk, A. Vissink, B.F.A.M. van der Laan, C.R. Leemans, H.P. Bijl, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, and J.A. Langendijk, “NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors.,” Radiother. Oncol. **105**(1), 101–6 (2012).

11 C.A. Murdoch-Kinch, H.M. Kim, K.A. Vineberg, J.A. Ship, and A. Eisbruch, “Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy.,” Int. J. Radiat. Oncol. Biol. Phys. **72**(2), 373–82 (2008).

12 J.O. Deasy, V. Moiseenko, L. Marks, K.S.C. Chao, J. Nam, and A. Eisbruch, “Radiotherapy dose-volume effects on salivary gland function.,” Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl), S58–63 (2010).

13 T. Dijkema, C.P.J. Raaijmakers, R.K. Ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, “Parotid gland function after radiotherapy: the combined michigan and utrecht experience.,” Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

14 F.Y. Feng, H.M. Kim, T.H. Lyden, M.J. Haxer, M. Feng, F.P. Worden, D.B. Chepeha, and A. Eisbruch, “Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures.,” Int. J. Radiat. Oncol. Biol. Phys. **68**(5), 1289–98 (2007).

15 G. Vlacich, D.E. Spratt, R. Diaz, J.G. Phillips, J. Crass, C.I. Li, Y. Shyr, and A.J. Cmelak, “Dose to the inferior pharyngeal constrictor predicts prolonged gastrostomy tube dependence with concurrent intensity-modulated radiation therapy and chemotherapy for locally-advanced head and neck cancer.,” Radiother. Oncol. **110**(3), 435–40 (2014).

16 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F.A.M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

17 S.A.S. Raktoe, H. Dehnad, C.P.J. Raaijmakers, W. Braunius, and C.H.J. Terhaard, "Origin of tumor recurrence after intensity modulated radiation therapy for oropharyngeal squamous cell carcinoma.," Int. J. Radiat. Oncol. Biol. Phys. **85**(1), 136–41 (2013).

18 A.K. Due, I.R. Vogelius, M.C. Aznar, S.M. Bentzen, A.K. Berthelsen, S.S. Korreman, A. Loft, C.A. Kristensen, and L. Specht, "Recurrences after intensity modulated radiotherapy for head and neck squamous cell carcinoma more likely to originate from regions with high baseline [18F]-FDG uptake.," Radiother. Oncol. **111**(3), 360-5 (2014).

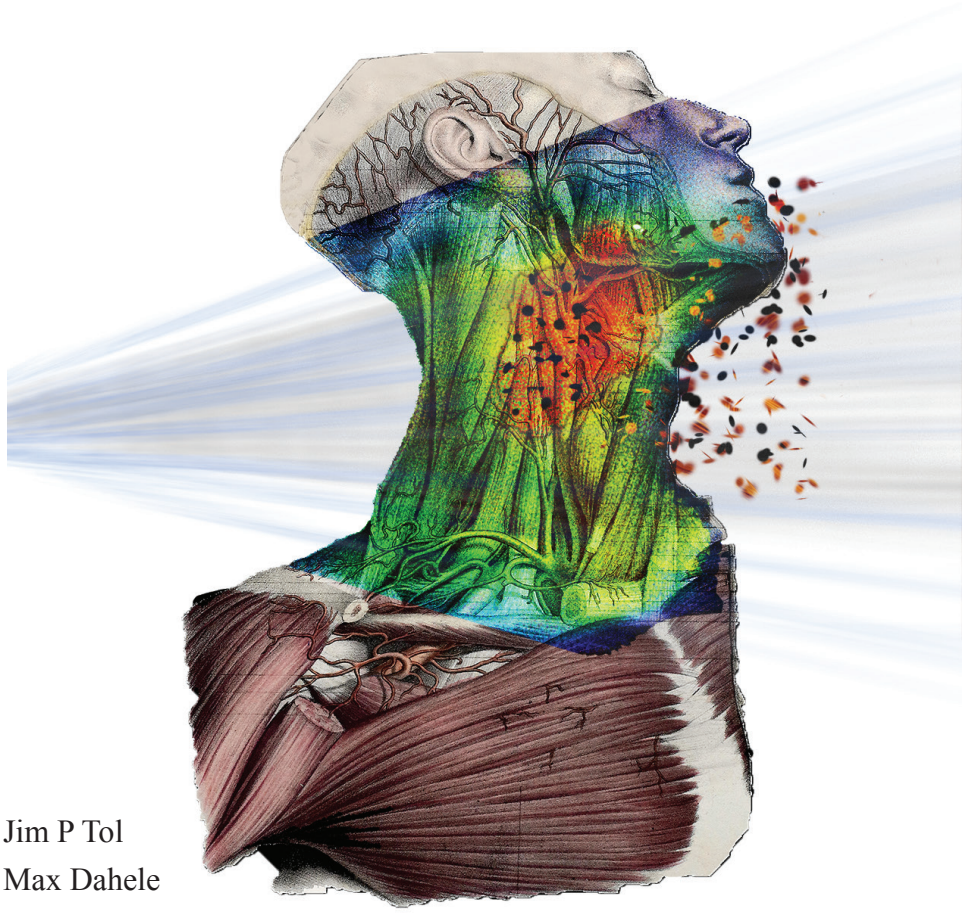
19 E. Bayman, R.J.D. Prestwich, R. Speight, L. Aspin, L. Garratt, S. Wilson, K.E. Dyker, and M. Sen, "Patterns of Failure after Intensity-modulated Radiotherapy in Head and Neck Squamous Cell Carcinoma using Compartmental Clinical Target Volume Delineation.," Clin. Oncol. (R. Coll. Radiol). **26**(10), 636-42 (2014).

20 C.M. Nutting, J.P. Morden, K.J. Harrington, T.G. Urbano, S.A. Bhide, C. Clark, E.A. Miles, A.B. Miah, K. Newbold, M. Tanay, F. Adab, S.J. Jefferies, C. Scrase, B.K. Yap, R.P. A'Hern, M.A. Sydenham, M. Emson, and E. Hall, "Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial.," Lancet Oncol. **12**(2), 127–36 (2011).

21 M.T. Guerrero Urbano, C.H. Clark, C. Kong, E. Miles, D.P. Dearnaley, K.J. Harrington, C.M. Nutting, and PARSPORT Trial Management Group, "Target volume definition for head and neck intensity modulated radiotherapy: pre-clinical evaluation of PARSPORT trial guidelines.," Clin. Oncol. (R. Coll. Radiol). **19**(8), 604–13 (2007).

Chapter 5

Automatic interactive optimization for volumetric modulated arc therapy planning



Jim P Tol
Max Dahele
Jarkko Peltola
Janne Nord
Ben J Slotman
Wilko FAR Verbakel

Radiation Oncology **10**(1), 75 (2015).

Abstract

Purpose

Intensity modulated radiotherapy treatment planning for sites with many different organs-at-risk (OARs) is complex and labor-intensive, making it hard to obtain consistent plan quality. With the aim of addressing this, we developed a program (automatic interactive optimizer, AIO) designed to automate the manual interactive process for the Eclipse treatment planning system. We describe AIO and present initial evaluation data.

Materials and Methods

Our current institutional volumetric modulated arc therapy (RapidArc) planning approach for head and neck tumors places 3-4 adjustable OAR optimization objectives along the dose-volume histogram (DVH) curve that is displayed in the optimization window. AIO scans this window and uses color-coding to differentiate between the DVH-lines, allowing it to automatically adjust the location of the optimization objectives frequently and in a more consistent fashion. We compared RapidArc AIO plans (using 9 optimization objectives per OAR) with the clinical plans of 10 patients, and evaluated optimal AIO settings. AIO consistency was tested by replanning a single patient 5 times.

Results

Average V95 & V107 of the boost planning target volume (PTV) and V95 of the elective PTV differed by ≤0.5%, while average elective PTV V107 improved by 1.5%. Averaged over all patients, AIO reduced mean doses to individual salivary structures by 0.9-1.6Gy and provided mean dose reductions of 5.6Gy and 3.9Gy to the composite swallowing structures and oral cavity, respectively. Re-running AIO five times, resulted in the aforementioned parameters differing by less than 3%.

Conclusions

Using the same planning strategy as manually optimized head and neck plans, AIO can automate the interactive Eclipse treatment planning process and deliver dosimetric improvements over existing clinical plans.

Introduction

An important goal in radiotherapy treatment planning is relatively homogenous irradiation of the planning target volumes (PTVs) whilst minimizing dose to nearby organs-at-risk (OARs). In many situations, especially when there is a complex OAR-PTV geometry, this can be achieved using some form of intensity modulated radiation therapy (IMRT), including volumetric modulated arc therapy (VMAT). Planning IMRT and VMAT treatments requires optimization of multileaf collimator (MLC) leaf positions to achieve suitable dose distributions. Different treatment planning systems (TPSs) offer different algorithms and interfaces to perform this optimization. For instance, some TPSs allow for interactive optimization, which involves presenting dose-volume histograms (DVHs) to the user and dynamically updating them while the user adapts specific dose-volume objectives during optimization. Ultimately, many factors influence the final treatment plan, including acceptance criteria for PTV dose coverage and homogeneity, optimization objectives and weightings for OARs and PTVs. If the specific TPS allows for interactive optimization, the experience of the planner, along with the interaction between the planner and the TPS, can also influence the obtained plan quality and may contribute to large variations between planners and centers[1,2].

Treatment planning has become increasingly complex over the years, particularly regarding the number of OARs that are included in the optimization. For example, radiotherapy treatments for head and neck cancer evolved from essentially contralateral parotid gland and spinal cord sparing[3] to include sparing of the ipsilateral parotid gland, the contralateral submandibular gland, multiple swallowing muscles and the oral cavity[4]. This further increases the difficulty of plan optimization and increases the likelihood that inconsistent planning results are obtained between planners. Automated planning techniques might assist in reducing such variation and allow for the creation of more consistent and high quality plans. Although automated planning is in its infancy, promising results have already been obtained using knowledge-based planning[5–13] and automated multicriteria plan optimization[14–16].

One of the most commonly used TPSs is Eclipse™ (Varian Medical Systems, Palo Alto, USA) which allows for interactive optimization of IMRT and VMAT plans. In our experience, when trying to create a plan that provides maximum OAR sparing for an individual patient (as opposed to creating plans where optimization stops once a pre-determined level of OAR sparing has been achieved, regardless of whether or not it could be improved upon for a given patient), the lowest achievable OAR doses and the trade-off between OAR sparing

and PTV dose homogeneity[17] are typically not known in advance of making the plan. This means that the optimal settings of the optimization objectives have to be determined during planning. This can be done during interactive optimization. As a result, interaction between the planner and the TPS is a key step in producing a good plan, while at the same time it presents a source of considerable variation in the manual planning process. We propose a novel approach to automate interactive VMAT planning using the Eclipse TPS that aims to address the increasing challenges of manual planning and attempts to meet the competing demands for consistent and high quality planning under conditions of increasing complexity, while limiting the total planning time. We believe that these challenges are common to many centers, especially where the creation of complex treatment plans is concerned. This report describes our automatic interactive optimizer solution (AIO) and presents an initial evaluation of its performance.

Materials and Methods

AIO was evaluated by creating simultaneous integrated boost plans for ten head and neck cancer patients that were previously treated using RapidArc. RapidArc is the VMAT approach of Varian Medical Systems, based on the work of Otto[18]. Prescribed doses were 54.25Gy to the elective PTV (PTV_E) and 70Gy to the boost PTV (PTV_B) in 35 fractions of 1.55Gy and 2Gy, respectively. A 5mm transition zone (PTV_T) was created between PTV_B and PTV_E to facilitate a dose fall-off between them. Table 1 shows the tumor site, stage and PTV volumes for each patient. For optimization and reporting purposes, a 6mm 'virtual' build-up region was used to obtain adequate target coverage in areas where the PTV approached the surface[19].

In all plans included in this study, optimization was performed using the progressive resolution optimizer (PRO) version 10.0.28 and followed by a 'continue previous optimization' (CPO) to improve PTV dose homogeneity[17]. Dose calculation was performed using the anisotropic analytical algorithm (AAA) version 10.0.28 with a 2.5mm grid size.

Table 1. Detailed information of the included head and neck cancer patients. The disease site, stage and volumes of the elective, transition and boost planning target volumes (PTV_E, PTV_T and PTV_B, respectively) and the volumes of the composite salivary and swallowing structures for all patients. The oral cavity was included as an OAR in 6 out of 10 patients.

Patient Number	Disease Site	Stage	PTV_E (cm^3)	PTV_T (cm^3)	PTV_B (cm^3)	$comp_{sal}$ [a] (cm^3)	$comp_{swal}$ [b] (cm^3)	oral cavity (cm^3)
1	Larynx	T2N2c	428.7	93.7	243.3	66.7	7.1	-
2	Oropharynx	T2N2b	441.7	48.3	188.6	82.3	16.5	-
3	Oropharynx	T2N2a	240.6	43.4	94.5	42.4	25.0	-
4	Oropharynx	T4N1	288.2	57.3	237.5	60.6	10.6	-
5	Oropharynx	T4aN1	280.8	79.5	164.5	78.9	32.0	36.7
6	Oropharynx	T4aN2b	288.2	57.3	237.5	60.6	10.6	14.6
7	Oropharynx	T4aN1	360.4	62.9	143.7	73.0	35.0	57.1
8	Oropharynx	T3N1	500.3	87.9	231.7	55.0	11.9	29.4
9	Oropharynx	T4aN2c	258.4	155.7	328.9	38.3	4.8	38.1
10	Hypopharynx	T2N3	264.0	137.0	607.0	70.8	10.3	60.6

[a] Composite salivary glands. Depending on degree of overlap with the PTVs and choice for inclusion by the treating clinician, $comp_{sal}$ could consist of some, or all, of the ipsilateral and contralateral parotid and submandibular glands.
[b] Composite swallowing muscles. Could consistent of some, or all, of the upper esophageal sphincter, upper and lower parts of the larynx, the superior, medial and inferior pharyngeal constrictor muscle, the cricopharyngeal muscle and the esophagus.

The Progressive Resolution Optimizer (PRO)

PRO is used to optimize the MLC apertures of arc fields in a treatment plan. A multi-resolution (MR) model is used, meaning that the angular dose representation starts with a crude approximation that gets finer as the optimization progresses[20]. Input parameters are the geometric characteristics of each field and a set of optimization objectives, see below, that can be adapted at any point during the optimization. The output of the optimizer is a control-point (cp) sequence, defining MLC configuration and MU count at each of the arc's 178 control points. Each structure's DVH-line is displayed and can be manipulated by adapting the optimization objectives, to attempt to meet clinical goals for PTV dose coverage and OAR doses. Each optimization objective has four input parameters: an optimization priority (P), a 2D-position on the DVH-graph representing the dose and volume goal ($dose_{goal}$, $volume_{goal}$), and information describing whether the $dose_{goal}$ is an upper (maximum) or lower (minimum) dose limit for the structure's DVH-curve.

The objective weighting ($objective_{weight}$) is derived from P using a heuristic power law-formula. A doubling of P results in a 32 time multiplication of $weight_{objective}$. To reduce

the occurrence of hot and cold spots, PRO also increases $objective_{weight}$ for objectives with $volume_{goal}$ values of 0% or 100%.

Each point (i) inside a structure, not fulfilling the objective, gets assigned a penalty cost. A structure with n points and m optimization objectives obtains a total cost value of:

$$(1) \qquad \frac{\sum_{j=1}^{m}\sum_{i=1}^{n} objective_{weight,j} \times (dose_i - dose_{goal,j})^2}{n}$$

The cost for an objective j is only taken into account for the range of voxels that violate the assigned dose-volume criteria. For example, if $dose_{goal}$ and $volume_{goal}$ are set to <5Gy and 10%, respectively, 10% of the structure volume may have doses higher than 5Gy without contributing to the cost function for j. If this criterion is violated, the cost function is calculated for the structure points with $dose_i$ values ranging from 5Gy to the dose achieved at 10% volume of the DVH-curve, see Figure 1. The total cost function value is calculated by summing equation (1) over all structures that are included in the PRO optimization.

During optimization, new, mutated MLC configurations that result in the largest decrease to the total cost function are successively added to the control points. In the first MR-level, the optimizer divides the arc into eleven 16-cp long sequences. The optimizer randomly selects 8 sequences and makes simultaneous changes to MLC configurations, while taking into account physical limitations such as MLC leaf speed, gantry speed and dose rate. Three subsequent increases in the MR-level increase the number while decreasing the size of these cp-sequences to 22 times 8-cps, 44 times 4-cps and finally 88 times 2 cps. Changes in positioning of the MLC leafs are dictated by set of 0-8 cp-sequences found to contribute most to reducing the total cost function. These changes are incorporated into the subsequent iterations. At the start of each iteration, the optimization objectives, which might have been changed, are sent to the PRO algorithm, allowing it to find more optimal MLC configurations to satisfy the set objectives. If the total cost function has converged, the optimization automatically proceeds to the next MR-level. The user can pause the automatic progression to subsequent MR-levels allowing the user more time to make changes to the dose-volume objectives. In the earlier MR levels, PRO is more flexible towards changes in the optimization goal during optimization, as it can make large changes in the MLC leaf configurations at each iteration. The possibility to adjust to changed dose-volume objectives gets smaller in later MR levels as the modified cp-sequence length decreases.

Optimization using dose-volume objectives

Due to the power-law nature of the objective weightings, structures with higher prioritized objectives will have substantially more influence on the resulting plan. In our clinical

planning protocol, described below, optimization objectives of PTV structures are assigned higher priorities values than OAR objectives. OAR objectives can therefore be gradually pushed to lower dose values at each iteration during optimization without significantly affecting the DVH's of the PTVs. This concept forms the basis of our institutional approach to manual interactive optimization and our automation of it, explained below.

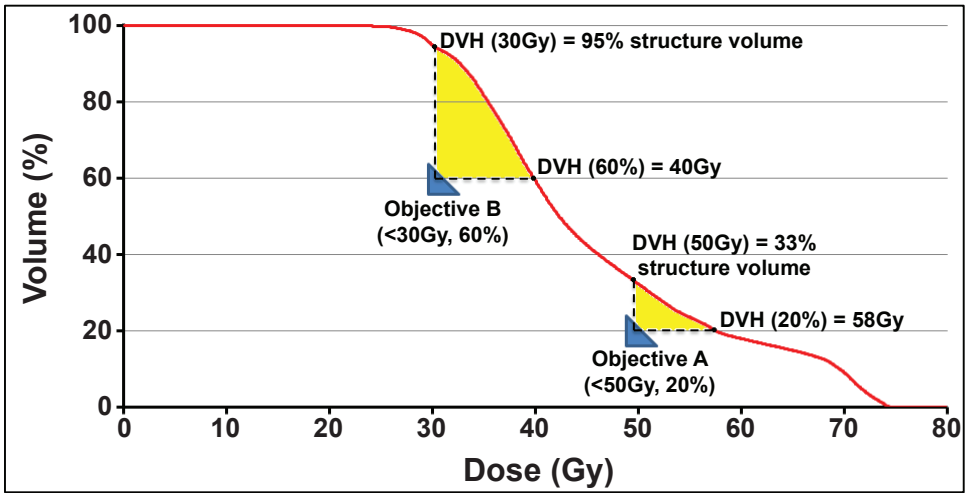


Figure 1. A schematic representation of the local force exerted on the dose-volume histogram (DVH)-line (yellow area) by two optimization objectives (blue triangles) Since structure points only contribute to an objective's cost function if the set dose-volume criteria are violated, the objectives can lower the DVH-curve locally. Objective A contributes 13% (33%-20%) of structure points to calculating the cost function, with $dose_i$ values ranging from 50-58Gy. Because objective B is placed further from the DVH-curve, the number of included structure points is larger (35%), with doses ranging from 30-40Gy. In case of overlap between the yellow areas, the included structure points would contribute to two separate cost functions.

Interactive optimization

PRO uses a simplified fast dose calculation algorithm to display DVH lines during the optimization process (Figure 2), visualizing changes in PTV and OAR doses while optimization of the MLC leaf positions is being performed. As previously explained, the priority and distance (i.e. dose difference) of an optimization objective to its corresponding DVH-line determines the 'effort' made by the optimization algorithm to meet this objective.

Initial evaluation of the algorithm during the clinical introduction of RapidArc at our department in 2008 resulted in the current institutional planning approach for head and neck tumors[21–23]. Because the patient-specific favorable dose-volume values of the OAR optimization objectives are typically not known at start of planning, this approach involves continuous user interaction, adapting the location of the OAR objectives so that they maintain a certain diagonal distance to their respective DVH-line while the optimization progresses through the different MR-levels. Parallel OAR, such as the parotid glands, submandibular glands and swallowing muscles, typically have 4-5 objectives, distributed along the volume axis of the DVH (prioritized at 80-90, but can be lowered for individual OARs after discussion between clinician and planner) of which 3-4 are interactively adjusted. Serial OARs such as the spinal cord and brainstem, have single maximum point dose objectives that remain unchanged during optimization. Objectives for the PTVs have priority values of 130 and remain unchanged during planning. Interactively adapting multiple objectives along a DVH results in equal effort / attention directed at sparing over the entire volume range and works well for parallel OAR in which the mean dose has been demonstrated to correlate with toxicity[24,25]. Since continuous pulling of the OAR objectives could lead to local underdosage of the PTVs near those OARs, parts of the PTV within 5 mm of the OARs are defined as an extra PTV volume with separate minimum dose objectives. Interactive optimization using this approach creates treatment plans with a high probability of satisfying our institutional guidelines for coverage and homogeneity of the PTV. The aims are that at least 99% of the boost volume is covered with 95% of the prescribed dose (V95 ≥ 99%) and no more than 4% should receive a dose greater than 107% (V107 ≤ 4%). Corresponding aims for the elective volume are V95 ≥ 98% and V107 ≤ 15%.

Head and neck cancer treatment plans currently include 2 PTVs planned to receive different doses, and up to 12 individual salivary and swallowing structures are routinely used. In addition, the oral cavity was included during optimization for 6 out of 10 patients. Because so many OARs are included, the continuous manual adjustment of multiple optimization objective locations along numerous DVH-lines has become increasingly challenging.

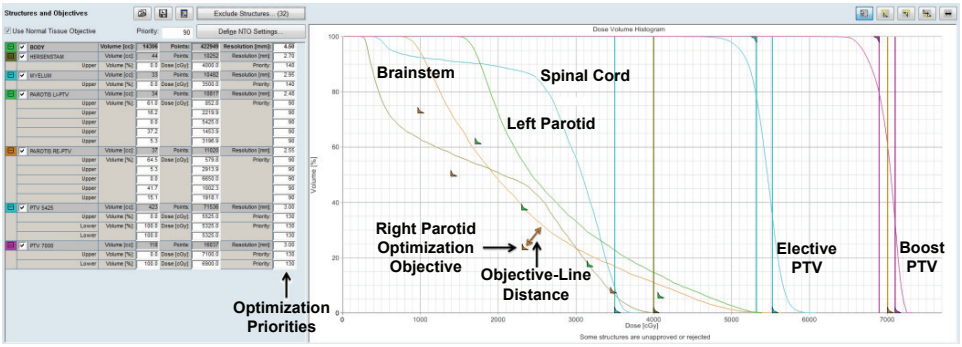


Figure 2. The Eclipse optimization window for a simple head and neck cancer patient showing dose-volume histogram (DVH)-lines and optimization objectives of the elective and boost PTVs, spinal cord and brainstem, and two parallel OARs (left and right parotids).

5

Automatic optimization solution

An automated alternative to manual interaction was developed to interactively optimize a large number of OARs consistently (automatic interactive optimizer, AIO). AIO was coded in Lazarus, a free Pascal compiler (*http://www.lazarus.freepascal.org/*). The flowchart in Figure 3 provides an overview of an automated optimization using AIO. Because Eclipse does not allow for modifications of the optimization algorithm and interface, AIO was designed to take over interactivity of the user by automatically moving the mouse cursor and adapting the position of the OAR objectives on the screen. If Eclipse would permit access to the optimization algorithm, this process could easily be implemented in the optimizer itself.

Before the PRO optimization is started, the user inputs the coordinates of the optimization window (step 1, Figure 3), AIO scans this window determining the initial location of all OAR objectives (step 2, Figure 3 and Figure 4A). Because color-coding is used to differentiate between the various OAR objectives, each OAR should be assigned a unique color. When the optimization is started (step 3, Figure 3), AIO scans the optimization window and determines the location of every DVH-line (step 4, Figure 3), based on the unique color of each OAR. AIO then places every OAR objective at a fixed diagonal pixel distance (the so-called 'objective-line' distance [OLD]) from the respective DVH-line (step 5, Figure 3 and Figure 4B & 4C). Because AIO moves the optimization objectives as soon as the DVH-lines appear on the screen, their initial locations do not influence the resulting plan quality. If the new objective position is already occupied with a different optimization objective, AIO searches for the next available position while keeping a minimum distance of 20 pixels between the objectives to

avoid overlap. To allow convergence of the cost function and let PRO advance to the next MR-levels in the optimization process, AIO is paused for T seconds (step 6, Figure 3), after which steps 3-5 are repeated until PRO completes the RapidArc optimization (step 7, Figure 3).

By automatically determining the position of the OAR DVH-lines and repositioning the optimization objectives, AIO removes the limitation that is posed by the relative knowledge, experience, precision and attention of the planner in the interactive planning process, allowing far more frequent and consistent updating of the OLD. If the OLD value is equal for each OAR, consistent positioning of the optimization objectives using AIO results in a constant cost function for each OAR objective throughout the optimization process. This allows PRO to balance sparing of the various OARs, depending on the assigned optimization priorities. The OLD and PTV / OAR optimization priorities can be adjusted in order to meet user-specific criteria for PTV dose coverage and homogeneity[26]. Additionally, because the location of every OAR objective is adjusted automatically, more optimization objectives can be used simultaneously throughout the optimization. Since the $dose_{goal}$ and $volume_{goal}$ values of these objectives determine the range of dose values taken into account when calculating the cost function (see Figure 1), distributing optimization objectives over the entire volume range may offer improved sparing compared to using 3-4 objectives per OAR.

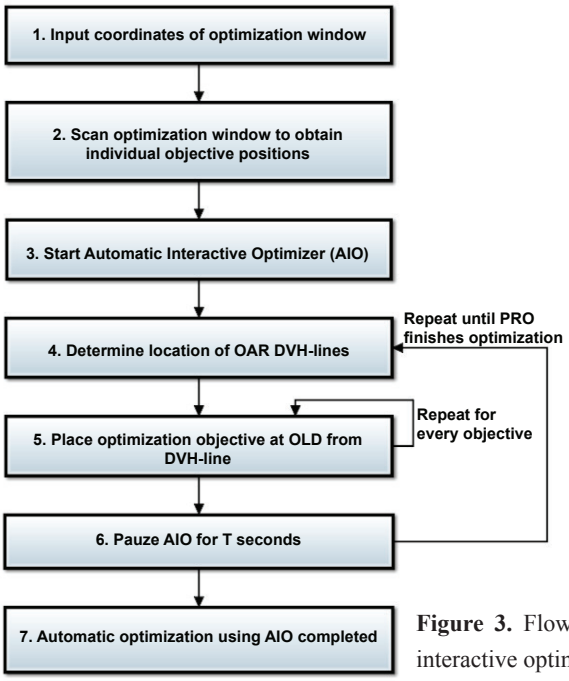


Figure 3. Flowchart of the automatic interactive optimizer (AIO).

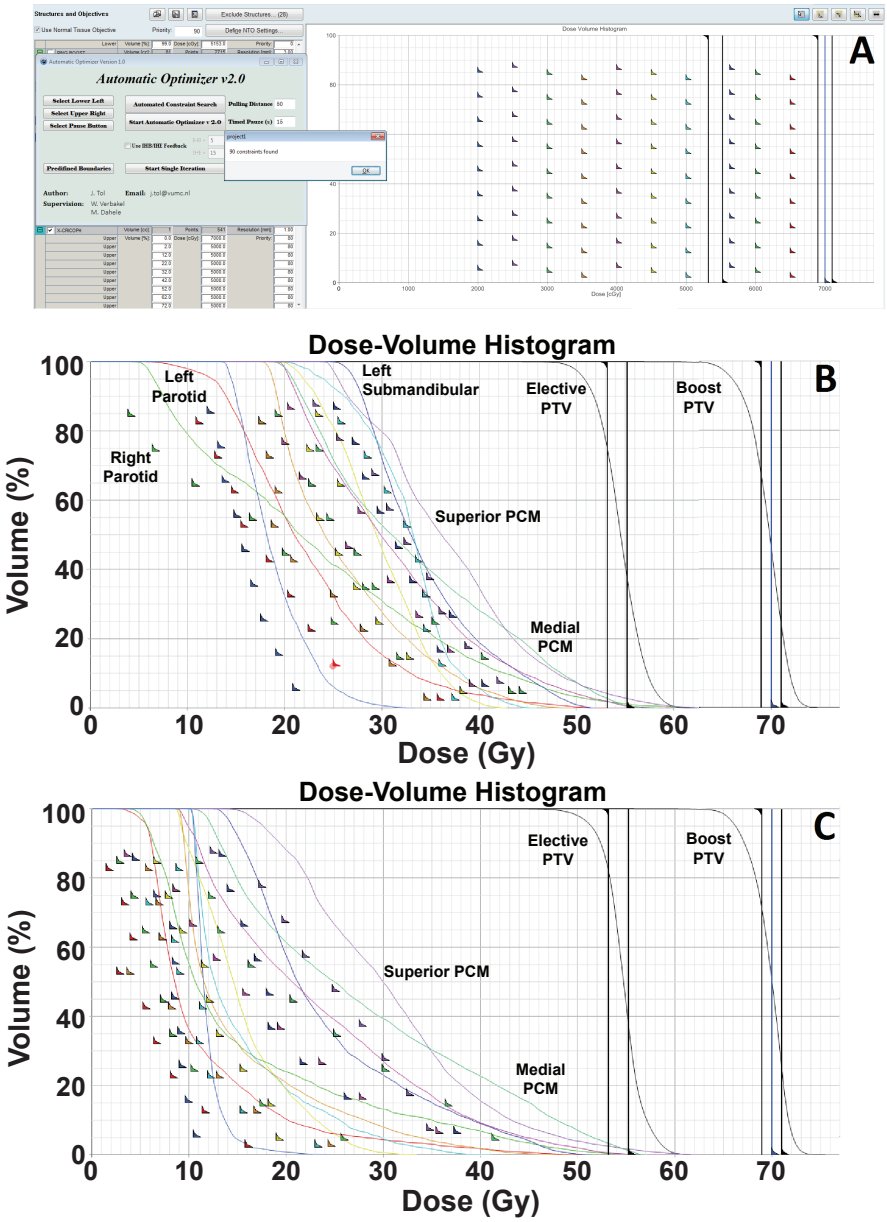


Figure 4. Different stages of optimization using the automatic interactive optimizer (AIO). Figure A) shows the initial position of all OAR optimization objectives along with the graphical user interface of AIO. Figure B) shows the optimization window after 3 minutes while AIO is being used for a complex head and neck case. Figure C) shows the optimization window after 11 minutes. In this case, the full optimization took approximately 29 minutes. Shown abbreviations: PTV = Planing target volume and PCM = Pharyngeal constrictor muscle.

Evaluation endpoints

For each patient, the manually optimized (clinical) plan is compared to the (nominal) AIO plan created using 9 different optimization objectives per OAR, an OLD value of 50 pixels and a pause time T of 15 seconds. The OLD value is screen-dependent, a value of 50 was chosen to reflect the OLD used clinically and corresponds to roughly 5% of the width of the optimization window at a screen resolution of 1980x1080 pixels. In the AIO plans, the optimization priorities of PTV and OAR objectives are kept equal to those used in the clinical plans. To investigate consistency of AIO plans, a single patient was re-planned 5 times using the same AIO settings. To determine the influence of the initial placement of the optimization objectives on the resulting plan, optimization of 5 patients was performed again without interactively repositioning the optimization objectives. The resulting optimization objectives in the nominal AIO plans were used as the initial location of the optimization objectives in these non-interactively optimized plans, and remained unchanged throughout the optimization.

To evaluate resulting plan quality when using different settings of the AIO, additional plans were created using (i) 3 and 6 planning objectives per OAR, (ii) OLD values of 25 and 75 pixels and (iii) pause times of 7.5 and 30 seconds while the other parameters were kept at the aforementioned values.

The following dosimetric parameters were evaluated after normalizing the AIO plans to deliver the same mean dose to PTV_B as the respective clinical plan: 1) Mean doses to the oral cavity (D_{OC}), individual and composite salivary glands (D_{sal}) and swallowing muscles (D_{swal}), 2) PTV_B and PTV_E volumes receiving less than 95% (V95) and more than 107% (V107) of the prescribed dose. The size of the oral cavity, composite salivary ($comp_{sal}$) and swallowing structures ($comp_{swal}$) is shown in Table 1. To determine whether improved sparing using AIO resulted in increased dose deposition in the remainder of the body, the mean dose to a body – PTV structure (obtained by subtracting the combined PTV from the body contour) was determined, along with the body – PTV volume receiving >5Gy (V5Gy), >30Gy (V30Gy) and >50Gy (V50Gy).

Results

AIO and clinical plan results were averaged over all 10 patients and are summarized in Table 2. Maximum spinal cord and brainstem doses were similar and considered clinically acceptable in all plans. Although PTV_B and PTV_E V95 values were marginally worse, consistently improved OAR sparing was obtained in the nominal AIO plan that

used an OLD of 50 pixels, 9 optimization objectives per OAR, and a pause time T of 15 seconds. Compared to the clinical plan, the nominal AIO plan reduced mean dose to individual salivary structures by on average 0.9-1.6Gy, and reduced D_{sal} from 26.2±5.5Gy to 25.0±5.8Gy. The oral cavity and individual swallowing muscles saw a larger dose reduction, with composite metrics D_{OC} / D_{swal} decreasing by 3.9Gy / 5.6Gy, on average. This increased OAR sparing using AIO came at the expensive of a 0.1% higher body – PTV mean dose on average, while V5Gy, V30Gy and V50Gy were respectively 0.2%, 0.4% and 0.4% higher. AIO plans required 2.8% more monitor units (MU), going from an average of 492MU per 2Gy fraction in the 10 clinical plans to 506, suggesting that additional MLC modulation was needed to achieve improved OAR sparing.

Figure 5 shows typically resulting DVH-lines and dose distributions of a nominal AIO plan (dashed lines) and clinical plan (solid lines) for patient 2. Using AIO, D_{sal} and D_{swal} decreased by 2.3Gy and 5.7Gy, respectively while approximately equal coverage of the boost and elective PTVs was obtained.

Running AIO five times on a single patient showed variations in D_{sal} and D_{swal} no greater than 0.9% and 2.0%, respectively. V95 of PTV_B and PTV_E ranged from 98.6-99.0% and 97.1-97.7% respectively, while V107 ranged from 0.5-1.2% and 14.3-17.3%. For this patient, the PTV coverage values were slightly lower than the institutional guidelines because the AIO plans were normalized to receive the same mean PTV_B dose as the original clinical plan. In clinical practice this would have been solved by either adjusting the normalization by a fraction of a percent, or by running an additional CPO calculation while slightly increasing the priority on PTV optimization objectives. The non-interactive optimization of the nominal AIO plans of 5 patients resulted in PTV_B / PTV_E V95% values within 0.6% / 0.3% of their respective nominal AIO plan, while D_{OC} / D_{sal} / D_{swal} varied no more than 0.3Gy / 0.4Gy / 0.8Gy.

Using 3 instead of 9 optimization objectives per OAR was associated with an increase in D_{OC}, D_{sal} and D_{swal} by 2.3Gy, 2.8Gy and 3.6Gy compared to the nominal AIO plan, respectively (Table 2), while PTV dose homogeneity improved. Running AIO with 6 objectives per OAR resulted in similar PTV coverage and homogeneity as the nominal plan, while D_{OC}, D_{sal} and D_{swal} were respectively 0.9Gy, 0.8Gy and 0.7Gy higher. Although an OLD of 25 pixels improved V95 / V107 of PTV_B and PTV_E by 0.6% / 0.7% and 1.0% / 3.5%, respectively, D_{OC} / D_{sal} / D_{swal} values increased by respectively 2.0Gy / 2.7Gy / 3.9Gy. The opposite was noticed using an OLD of 75 pixels, which degraded V95 / V107 of PTV_B and PTV_E by 1.4% / 2.2% and 1.7% / 3.5%, respectively, while D_{OC} / D_{sal} / D_{swal} decreased by 1.1Gy / 0.8Gy / 1.4Gy.

Table 2. Automatic Interactive Optimizer (AIO) and clinical plan results. AIO and clinical plan results averaged over all 5 patients using different AIO settings. Clinically acceptable doses to the brainstem and spinal cord were achieved in all plans.

Plan	Clinical	Nominal AIO	AIO OLD=25[d]	AIO OLD=75[d]	AIO 3 objs/OAR[e]	AIO 6 objs/OAR[e]	AIO T=7.5s[f]	AIO T=30s[f]
Boost PTV (%)								
V95%[a]	99.1 ± 0.2	98.9 ± 0.5	99.5 ± 0.2	97.5 ± 1.2	99.2 ± 0.5	98.9 ± 0.5	98.7 ± 0.6	98.5 ± 0.9
V107%[a]	1.1 ± 1.4	1.2 ± 1.6	0.5 ± 0.6	3.4 ± 3.6	0.5 ± 0.7	1.1 ± 1.2	1.9 ± 1.9	1.6 ± 1.6
Elective PTV (%)								
V95%	98.2 ± 1.1	97.7 ± 0.7	98.7 ± 0.6	96.0 ± 1.6	98.3 ± 1.0	97.8 ± 1.0	97.4 ± 1.0	97.3 ± 1.0
V107%	14.5 ± 3.8	13.0 ± 5.4	9.5 ± 4.6	16.5 ± 6.2	10.3 ± 4.7	12.2 ± 5.0	13.7 ± 5.6	13.6 ± 5.1
Max Dose (Gy)								
Spinal Cord	39.0 ± 3.7	39.1 ± 4.1	38.3 ± 4.0	39.8 ± 3.7	38.5 ± 4.2	38.8 ± 4.0	39.0 ± 3.8	38.3 ± 4.9
Brainstem	37.7 ± 7.7	37.3 ± 8.0	37.7 ± 8.3	38.8 ± 7.8	37.0 ± 7.2	37.6 ± 8.3	37.0 ± 7.9	37.5 ± 8.3
Mean Dose (Gy)								
CL parotid	19.6 ± 5.6	18.7 ± 5.7	21.4 ± 6.9	17.6 ± 5.3	21.4 ± 6.7	19.5 ± 6.0	18.5 ± 5.5	19.0 ± 5.9
IL Parotid	31.2 ± 8.1	29.9 ± 8.4	32.0 ± 8.5	27.5 ± 7.1	32.1 ± 8.7	30.5 ± 8.3	29.5 ± 7.9	29.8 ± 8.1
CL SMG[b]	31.2 ± 7.4	29.6 ± 8.0	34.0 ± 8.6	27.7 ± 7.3	33.9 ± 8.6	30.6 ± 8.3	29.4 ± 7.0	30.5 ± 7.4
Oral Cavity	28.6 ± 7.0	24.7 ± 7.9	26.7 ± 8.7	23.6 ± 7.6	27.0 ± 10.3	25.6 ± 8.3	24.2 ± 6.9	26.1 ± 7.4
$Comp_{sal}$[c]	26.2 ± 5.5	25.0 ± 5.8	27.7 ± 6.5	24.2 ± 5.3	27.8 ± 6.2	25.8 ± 5.9	24.7 ± 5.4	25.2 ± 5.7
$Comp_{swal}$[c]	28.7 ± 8.0	23.1 ± 8.7	27.0 ± 9.5	21.7 ± 8.7	26.7 ± 9.5	23.8 ± 9.2	22.5 ± 8.8	23.4 ± 8.7

[a] PTV volumes receiving ≥95% and ≤107% of the prescribed dose

[b] Contralateral submandibular gland

[c] Composite salivary and swallowing structures

[d] AIO plans with: objective-line distance (OLD) of 25 / 75 pixels, 15 second pause time seconds, 9 objectives per OAR

[e] AIO plans using 3 or 6 objectives (objs) per OAR, a pause time of 15 seconds and an OLD of 50 pixels

[f] AIO plans using pause times (T) of 7.5 or 30 seconds, 9 objectives per OAR and an OLD of 50 pixels.

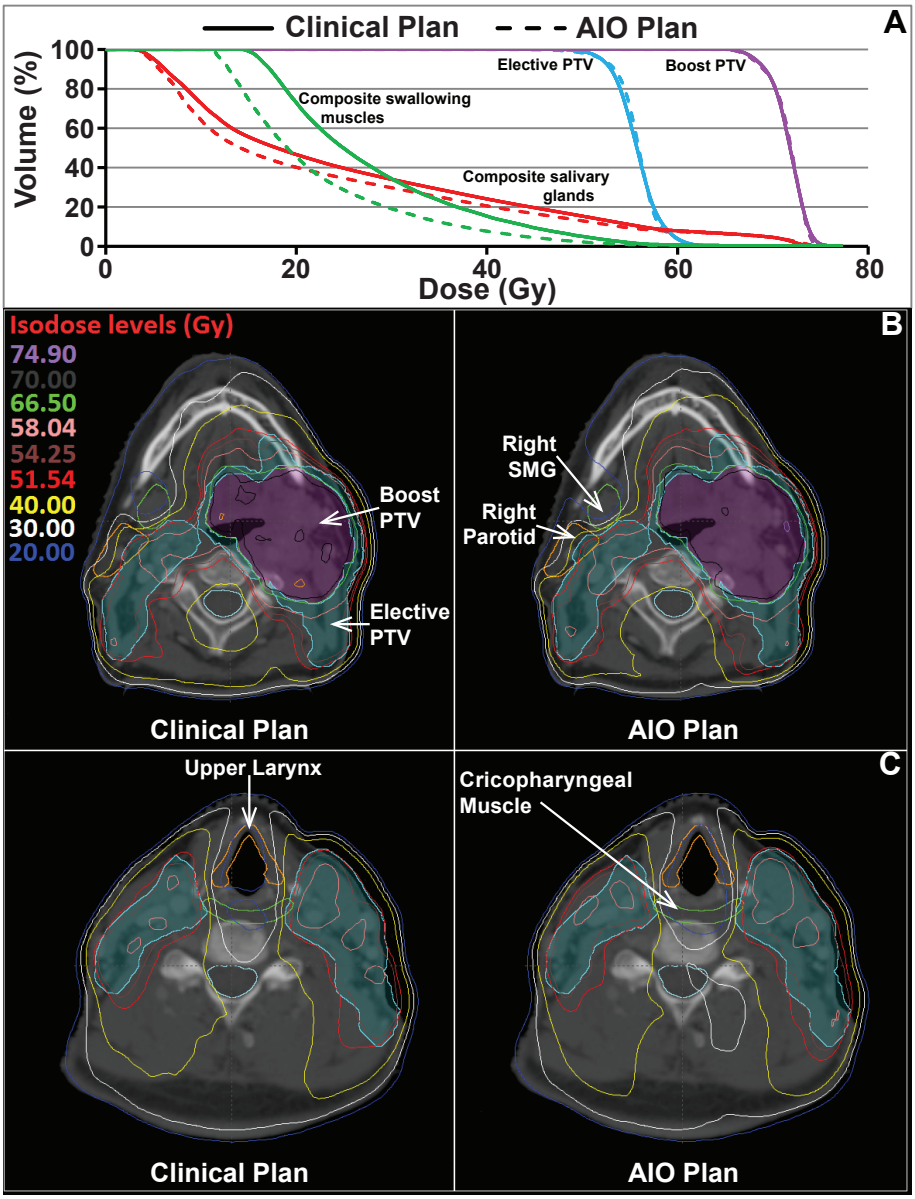


Figure 5. A) shows the dose-volume histogram (DVH)-lines of an AIO plan (dashed lines) and clinical plan (solid lines) for patient #2. Mean dose to the composite salivary glands (red) and swallowing muscles (green) decreased while approximately equal coverage of the boost and elective PTVs (purple and blue, respectively) was obtained. B) shows the dose distributions on the level of the salivary glands while C) shows the dose distributions on the level of the swallowing muscles.

Changing pause times of the AIO had little impact on the PTV coverage / homogeneity and OAR sparing. On average, AIO plans with a T = 7.5s achieved slightly lower OAR doses relative to T = 15s, while T = 30s increased them. This increase was associated with a decrease in average optimization time, which was approximately 56, 32 and 16 minutes for T values of 7.5, 15 and 30 seconds using an Intel® Xeon® E2520 2.40Ghz CPU in combination with a distributed calculation grid. Although the magnitude was influenced by the complexity of the patient, the number AIO iterations (steps 4 to 6, Figure 3) decreased for increasing T values. For patient 1 for example, AIO performed 85, 50 and 15 iterations using pause times of 7.5, 15 and 30 seconds, respectively.

Discussion

This study presented an initial evaluation of our in-house developed automatic interactive optimizer (AIO). AIO was designed to automate our interactive optimization protocol for head and neck cancer using the Eclipse treatment planning system in order to improve plan quality and consistency. Under the planning constraints posed during optimization by the minimum / maximum dose objectives on the PTVs and single point dose objectives on serial OARs (e.g. spinal cord and brainstem), AIO gradually positions the optimization objectives of parallel OARs (e.g. salivary glands, swallowing muscles and oral cavity) at lower dose values throughout the optimization process. Compared to the clinical plans, nominal AIO plans achieved substantially improved sparing of the swallowing muscles and oral cavity, although a small decrease in PTV dose homogeneity was found in some patients, which was considered clinically acceptable[17,26]. On the other hand, if it is desired, then with minimal loss in OAR sparing, the PTV homogeneity can be improved after the AIO optimization by running a relatively short (non-interactive) CPO with higher priorities on the PTV optimization objectives. It is important to state that the use of AIO does not result in any changes in the quality assurance process. Even though AIO resulted in dosimetric improvements over clinical plans, the final DVHs and dose distributions should still be routinely evaluated by the planning dosimetrist, radiation oncologist and physicist, prior to acceptance for treatment. This is no different from treatment plans created manually or using different automatic planning approaches.

Once AIO is running, it is fully automated, thereby allowing the planner to invest time in other activities while the optimization is being performed. Running AIO optimizations in a virtual machine (Oracle VirtualBox; *http://www.virtualbox.org*) allows the user to retain control of their cursor. They can therefore continue using their PC while AIO is running. A virtual machine also allows a custom screen resolution to be defined. Larger resolutions

may prevent potential overlap of the optimization objectives during optimization, which leads to suboptimal placement.

The effect of changing several user-definable AIO parameters was investigated and showed: 1) The selected objective-line distance (OLD) influences the trade-off between PTV dose homogeneity and OAR sparing. These results can be expected because in PRO, increasing or decreasing the OLD will relatively increase or decrease the cost of these OAR optimization objectives, compared to the PTV objectives. 2) An increased number of optimization objectives generally improved OAR sparing because a reduction in OAR doses was attempted at a greater range of structure volumes. 3) The desired balance between optimization time and resulting plan quality is determined by the chosen value of T. Larger values for T resulted in less AIO iterations and in effect, the optimization objectives were not placed at a constant distance from their DVH-line at all times. Lower values of T increase the number of AIO iterations and optimization time, allowing the optimizer in theory to minimize the cost function better. However, no marked improvements were noted using a T value of 7.5s instead of 15s in this study.

Nominal AIO plans, generated with an OLD of 50 pixels, 9 optimization objectives per OAR and a T value of 15 seconds, provided a good balance between optimization time, OAR sparing and PTV dose homogeneity, and would therefore provide a reasonable starting point for treatment planning. The present results can also be generalized to manual interactive planning in the Eclipse TPS. The initial 'grid' of optimization objectives that was used (Figure 4A) for example, indicates that AIO does not need a proper initial prediction of the objective starting locations to obtain high quality plans, as long as the optimization objectives are repositioned when the DVH-lines appear in the optimization window. This assumption was validated by re-optimizing the nominal AIO plans of 5 patients non-interactively, using the final placement of the interactively determined optimization objectives. The resulting plans provided similar OAR sparing to the interactively optimized AIO plans.

AIO was solely designed to automate interactive optimization in the Eclipse TPS. However, we do not perceive this as a weakness, as many automated planning solutions will ultimately be vendor or TPS specific. However, since VMAT optimization algorithms generally start with coarse MLC configurations that get progressively finer as the optimization progresses in one way or another, the core concept of AIO, pushing OARs to lower doses throughout the optimization process while the MLC apertures are being defined, may also be transferable to different TPS's. Although AIO was only used in combination with our clinical OAR

and PTV optimization priorities so that it generated plans fulfilling our clinical plan acceptance criteria, we expect it to work equally well for other criteria by adjusting the optimization priorities for PTVs and the OLD. Running the same plan several times on the same patient, AIO demonstrated a satisfactory degree of consistency in OAR sparing and PTV dose homogeneity, demonstrating the reproducibility of the automated optimization approach[26]. Because of the consistent placement of the optimization objectives throughout the optimization, we expect that AIO optimized plans are well suited for treatment planning comparisons and planning studies where generally the user dependency of setting and interacting with parameters can substantially influence plan results[1].

Some potential limitations deserve comment. The present study focused on the evaluation of AIO for VMAT head and neck cancer planning. This was because VMAT is our standard delivery method for intensity modulation and head and neck is one of the sites where planners have to routinely deal with many different OARs during interactive optimization. We did not use the latest version of PRO, although the version used in this study (v10.0.28) is the one that is clinically used in our department. Preliminary tests indicate that AIO can also be used with newer versions of PRO (11.0.31 to 13.5.10). The current analysis investigated replacing manual interaction with software-controlled interaction for a particular planning strategy in which constant OLD are used and priorities are held constant. This was done to replicate our current institutional planning strategy. This strategy also requires that the user knows the input priorities that generally result in acceptable plans for the majority of patients. We have used a standard set of priorities for PTV and OAR, derived from our clinical practice, and we have not attempted to investigate the influence of applying different priorities. Although AIO has been specifically designed to reduce mean doses of OARs, other planning strategies (e.g. focused on high or low dose reduction) can also be incorporated. AIO was designed to automate an important step in planning, it does not automate the complete planning process. Optimality of a plan may depend on more than the optimization alone: the number of (partial) arcs[27], chosen collimator angles, isocenter positions and jaw settings, all which have to be input by the user, may also influence the achieved plan dosimetry. In addition, we recognize that there may be selected cases where planning teams may wish to use manual interaction. The use of AIO does not prohibit switching to manual optimization if desired.

Another automated planning solution described in the literature is automated multi-criteria optimization, typically using a 'wish list' that contains PTV and OAR optimization priorities and weights. For every OAR, weightings are successively increased until either the required OAR sparing is achieved or the specified PTV dose

homogeneity / sparing of a higher prioritized OAR is compromised[14–16]. The downside of this approach could be that it prevents reduction of dose to low prioritized OAR, even though a substantial reduction could be obtained with only a minor increase in dose to a higher prioritized OAR. This does not occur when using AIO because, apart from the optimization objectives that are assigned, all the optimization objectives are continuously pushed to lower doses. This is illustrated in a patient where AIO decreased the mean dose of three different swallowing muscles by >5.5Gy, while increasing mean dose to the contralateral submandibular gland by 1.8Gy. Although the clinician should decide whether this is trade-off is deemed acceptable, it illustrates that a small increase in dose to a single OAR may lead to a larger dose reduction in multiple OARs.

The present report described the development and initial evaluation of AIO, automating and improving the interactive optimization process in the Eclipse TPS. AIO consistently allows the user to more fully exploit the potential of interactive optimization in Eclipse. Importantly, this intermediate interface requires no modification of the TPS. Preliminary testing shows that AIO can deliver dosimetric improvements over manually optimized head and neck plans. AIO has been clinically implemented for head and neck treatment planning.

References

1 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

2 I.J. Das, C.-W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

3 A. Eisbruch, L.A. Dawson, H.M. Kim, C.R. Bradford, J.E. Terrell, D.B. Chepeha, T.N. Teknos, Y. Anzai, L.H. Marsh, M.K. Martel, R.K. Ten Haken, G.T. Wolf, and J.A. Ship, "Conformal and intensity modulated irradiation of head and neck cancer: the potential for improved target irradiation, salivary gland function, and quality of life.," Acta Otorhinolaryngol. Belg. **53**(3), 271–5 (1999).

4 H.P. van der Laan, A. Gawryszuk, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, E.W. Korevaar, O. Chouvalova, K. Wopken, H.P. Bijl, and J.A. Langendijk, "Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction.," Radiother. Oncol. **107**(3), 282–7 (2013).

5 E.M. Quan, J.Y. Chang, Z. Liao, T. Xia, Z. Yuan, H. Liu, X. Li, C.A. Wages, R. Mohan, and X. Zhang, "Automated volumetric modulated Arc therapy treatment planning for stage III lung cancer: how does it compare with intensity-modulated radio therapy?," Int. J. Radiat. Oncol. Biol. Phys. **84**(1), e69–76 (2012).

6 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **79**(4), 1241–7 (2011).

7 B. Wu, D. Pang, P. Simari, R. Taylor, G. Sanguineti, and T. McNutt, "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study.," Med. Phys. **40**(2), 021714 (2013).

8 X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

9 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

10 K.L. Moore, R.S. Brame, D.A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **81**(2), 545–51 (2011).

11 J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Evaluation of a knowledge-based planning solution for head and neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **91**(3), 612–20 (2015).

12 A. Fogliata, F. Belosi, A. Clivio, P. Navarria, G. Nicolini, M. Scorsetti, E. Vanetti, and L. Cozzi, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer.," Radiother. Oncol. **113**(3), 385–91 (2014).

13 A. Fogliata, P.M. Wang, F. Belosi, A. Clivio, G. Nicolini, E. Vanetti, and L. Cozzi, "Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer.," Radiat. Oncol. **9**(1), 236 (2014).

14 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

15 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

16 D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld, "Multicriteria VMAT optimization.," Med. Phys. **39**(2), 686–96 (2012).

17 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

18 K. Otto, "Volumetric modulated arc therapy: IMRT in a single gantry arc.," Med. Phys. **35**(1), 310–7 (2008).

19 C. Thilmann, A. Zabel, S. Nill, B. Rhein, A. Hoess, P. Haering, S. Milke-Zabel, W. Harms, W. Schlegel, M. Wannenmacher, and J. Debus, "Intensity-modulated radiotherapy of the female breast.," Med. Dosim. **27**(2), 79–90 (2002).

20 E. Vanetti, G. Nicolini, J. Nord, J. Peltola, A. Clivio, A. Fogliata, and L. Cozzi, "On the role of the optimization algorithm of RapidArc(®) volumetric modulated arc therapy on plan quality and efficiency.," Med. Phys. **38**(11), 5844–56 (2011).

21 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–9 (2009).

22 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

23 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

24 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F. a M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

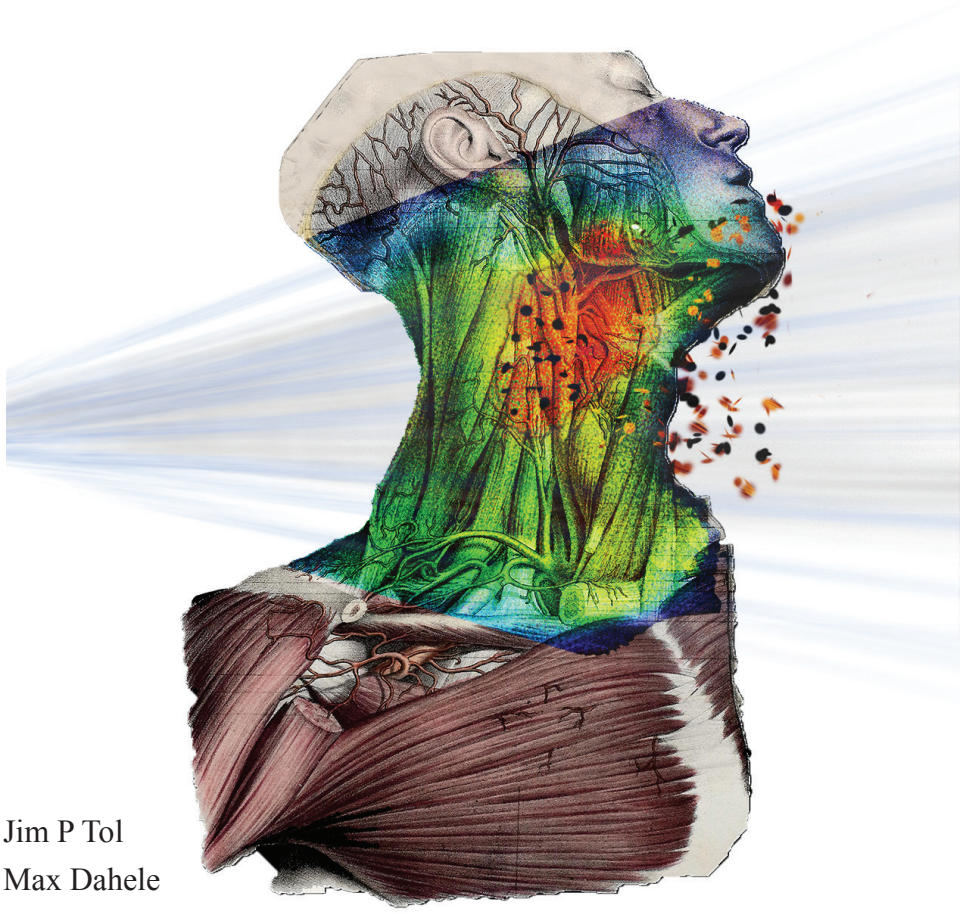
25 T. Dijkema, C.P.J. Raaijmakers, R.K. Ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

26 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel. "Different treatment planning protocols can lead to large differences in organ at risk sparing.," Radiother. Oncol. **113**(2), 267-271 (2014).

27 J.P. Tol, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Increasing the number of arcs improves head and neck volumetric modulated arc therapy plans.," Acta Oncol. **54**(2), 283–7 (2015).

Chapter 6

Detailed evaluation of an automated approach to interactive optimization for volumetric modulated arc therapy plans



Jim P Tol
Max Dahele
Alexander R Delaney
Patricia Doornaert
Ben J Slotman
Wilko FAR Verbakel

Medical Physics **43**(4), 1817-1829 (2016).

Abstract

Purpose

Interactive optimization during treatment planning requires intermittent adjustment of organ-at-risk (OAR) objectives relative to the dose-volume histogram (DVH)-line. This is a labor intensive process and the resulting plans are prone to variations in quality. Our in-house developed approach to automated interactive optimization (AIO) automatically moves the mouse cursor to adjust the position of the on-screen optimization objectives. This allows for the use of more objectives per OAR and a more frequent and consistent adjustment of these objectives during optimization. We report a detailed evaluation of AIO performance in support of its implementation for routine head and neck cancer (HNC) planning and an evaluation for locally advanced lung cancer (LC) planning which requires a different optimization strategy.

Materials and Methods

Volumetric modulated arc therapy (VMAT) AIO plans (APs) were created for seventy HNC patients with a simultaneously integrated boost and twenty LC patients, and benchmarked against their respective manually interactively optimized counterparts (MPs). The same set of optimization objectives and priorities was used for all APs, although planning target volume (PTV) optimization priorities could be increased manually in a subsequent 'continue previous optimization' calculation. HNC plans were benchmarked using mean dose to individual and composite OARs and elective / boost PTV (PTV_E / PTV_B) volumes receiving 95% and 107% of the prescription dose (V95% and V107%, respectively). A clinician performed blinded comparison of 20 APs and respective MPs. LC plans were compared using PTV V95% / V107%, contralateral lung (CL) volume receiving 5Gy (V5Gy), total lung – PTV (TL) V5Gy / V20Gy and esophagus and heart V40Gy / V60Gy / mean doses.

Results

For HNC, statistically significant improvements in sparing of all OARs, except for the ipsilateral submandibular gland and trachea, were obtained in the APs compared to MPs. Average mean dose to oral cavity, composite salivary and swallowing structures was 25.4Gy / 23.8Gy, 24.2Gy / 23.2Gy and 29.5Gy / 25.5Gy, respectively, for the MPs / APs. PTV heterogeneity was similar: in the APs, PTV_B V95% was 0.2% higher while PTV_B / PTV_E V107% was 0.4% / 1.0% lower. In 19 out of 20 HNC patients, the clinician preferred the AP, mainly because of better OAR sparing and PTV dose homogeneity. For LC, APs had a significantly lower CL V5Gy (6.1%), heart mean dose / V60Gy (0.9Gy / 1.2%) and

esophagus mean dose / V60Gy (0.9Gy / 2.8%), a non-significantly higher TL V20Gy (1.4%) and a slight, but significantly higher dose deposition to the body. PTV dose coverage and homogeneity were similar in the APs and MPs. AIO was considered sufficiently robust for clinical use in LC.

Conclusions

HNC and LC APs were at least as good as, and often of improved quality over MPs. To date, AIO has been clinically implemented for HNC planning.

Introduction

Treatment planning for intensity modulated radiotherapy (IMRT), or its rotational variant volumetric modulated arc therapy (VMAT), requires optimization of the multileaf collimator (MLC) leaf positions in order to try and fulfill specified dose-volume objectives. Different treatment planning systems (TPSs) use a varying range of algorithms and interfaces to achieve this. Some allow the planner to interactively adapt planning objectives during optimization[1], whereas others do not, and instead rely on an iterative approach by the planner to find optimal settings for the optimization objectives. Differences in planner-TPS interaction and planners' experience and skill in using the specific TPS are important sources of variation in plan quality[2–6]. This is exacerbated by increases in planning complexity, including large numbers of structures, faster optimization and increasing demands for organ-sparing dose distributions. A more automated approach to treatment planning reduces human-machine interaction, helping to address the increasingly demanding planning environment, and may improve planning consistency and quality. Different (semi) automatic solutions have been reported, such as knowledge-based planning and multicriteria optimization[7–11], and although these have obtained promising initial results[12–18], there is limited data concerning the combination of more complex disease sites and VMAT[21,22].

In a recent publication[21], we focused on the technical development of an in-house developed approach to automatic interactive optimization (AIO) for RapidArc® VMAT planning in the Eclipse™ treatment planning system (Varian Medical Systems, Palo Alto, USA) for which we investigated the influence of different parameter settings. In support of the clinical implementation of this approach a larger, more detailed investigation of AIO plan (AP) quality for complex head and neck cancer (HNC) treatments was required. Here, we report the results of this investigation, which includes a dosimetric comparison with clinical plans for 70 patients and a blinded clinical assessment for 20 of those plans. To further assess the potential of AIO, we also compared APs with manually interactively optimized plans (MPs) for 20 locally advanced lung cancer (LC) patients. Treatment planning for these patients demands a different approach to interactive optimization, and therefore tests the versatility of AIO.

Materials and Methods

Automatic interactive optimizer (AIO)

AIO automates the interactive optimization process by taking over control of the computer-screen cursor[21]. When the RapidArc optimization starts, AIO regularly scans the optimization window and determines the location of each DVH-line. AIO then automatically places every OAR objective at a specified distance below the DVH-line (within the confines of the user-specified boundaries of the optimization window). AIO ensures that an OAR is not neglected during optimization and allows for the inclusion of more objectives per OAR. While these objectives are being adjusted, the progressive resolution optimizer (PRO) optimizer is briefly and intermittently paused as part of the AIO process. Because the goal of AIO was to develop an automated planning solution that produces good plans in a single full optimization run for the vast majority of patients, standard sets of optimization objectives and priorities were used for all HNC patients (Table 1) and LC patients (Table 2) included in this analysis. This was in contrast to the MPs, where the planner could decide to adjust the number of optimization objectives and their priorities if this was considered appropriate. The spinal cord and brainstem were assigned fixed maximum dose objectives with a high priority that remained unchanged by AIO throughout the optimization. All plans included in this analysis were optimized using the progressive resolution optimizer (PRO) version 10.0.28, with the volume dose calculated using the anisotropic analytical algorithm (AAA) version 10.0.28 with a 2.5mm dose grid.

After optimization and dose calculation, a 'continue previous optimization' (CPO) was performed to improve the PTV dose homogeneity of all MPs and APs[22]. This is a brief optimization performed in only the last (4^{th}) multi-resolution level of the RapidArc optimization algorithm, without user interaction, and was developed to compensate for differences between final dose calculation with AAA and the simplified dose calculation algorithm used during optimization. This CPO typically has the largest effect on the low density regions of the PTV and it generally leads to a more homogeneous dose distribution in the PTV[22]. As increased OAR sparing in the APs could lead to underdosing of the PTVs near OARs, the user could decide, after reviewing the initial AP, to perform an additional CPO using higher PTV optimization priorities if the PTV dose coverage or homogeneity values were not clinically acceptable according to the institutional guidelines specified below). To allow for a fair comparison between PTV dose coverage, homogeneity and OAR sparing, all APs were normalized to deliver the same mean dose to the primary PTV as the respective MP.

Table 1. The standard set of optimization objectives used in this analysis to evaluate the performance of the automatic interactive optimizer (AIO) for head and neck cancer planning.

Structure	Optimization Objective	Priority	Dose/Volume Values	Dose-Difference[e]
Boost PTV[a]	2x Minimum	130	≥ 69Gy	-
	1x Maximum[d]	130	≤ 71Gy	-
Elective PTV[a]	2x Minimum	130	≥ 53.25Gy	-
	1x Maximum	130	≤ 55.25Gy	-
Transition PTV[a]	1x Minimum	130	≥ 54.25Gy	-
	1x Maximum	130	≤ 70Gy	-
OAR[b] overlap with Boost PTV	1x Minimum	130	≥ 69Gy	-
OAR[b] overlap with Elective PTV	1x Minimum	130	≥ 53.25Gy	-
Shoulders	1x Maximum	110	≤ 5Gy	-
Spinal Cord (+3mm PRV[c])	1x Maximum	120	≤ 37 (39)Gy	-
Brainstem (+3mm PRV[c])	1x Maximum	120	≤ 37 (39)Gy	-
Optimization objectives moved by AIO				
Parotid Glands	10x Upper	90	5 – 85%	3Gy
Submandibular Glands	10x Upper	85	5 – 85%	3Gy
Oral Cavity	10x Upper	80	5 – 85%	3Gy
Individual Swallowing Muscles[f]	10x Upper	80	5 – 85%	3Gy

[a] PTV = Planning Target Volume

[b] OAR = Organ-at-Risk

[c] PRV = Planning-at-Risk Volume

[d] Dose-volume objective at 0% volume

[e] Dose difference between the optimization objective and the respective DVH-line

[f] The same optimization objective settings were used for all individual swallowing muscles.

Head and neck cancer plans

Clinical plans of 70 HNC patients were used to benchmark AIO performance. The approach to planning has been described in detail previously[22]. In brief, dose prescription was set at 54.25Gy at 1.55Gy per fraction to the elective target volume (PTV_E) and 70Gy at 2Gy per fraction to the boost region (PTV_B) delivered as a simultaneous integrated boost (SIB). PTV_B consisted of an expansion of the clinical target volume, defined as a 5mm expansion of the gross tumor volume and biopsy-proven positive lymph node regions, and was edited for anatomic boundaries. PTV_E consisted of a similar expansion of the elective nodal regions, minus a 5mm transition zone (PTV_T) created to facilitate a steep dose fall-off between the PTVs. PTV expansions were chosen as 4mm above the clavicles, and 5mm below.

The reason for using different PTV expansions in our department was to account for a larger uncertainty in patient positioning below the shoulders. During the daily image-guided setup prior to treatment, the registration has the largest focus on the PTV_B region, which is most often lying in the cranial part of the entire PTV. Since rotations and deformations are not uncommon, the setup uncertainties can thus be larger at the caudal part of the PTV, for which we apply larger PTV expansions.

Table 2. The standard set of optimization objectives used in this analysis to evaluate the performance of the automatic interactive optimizer (AIO) for locally advanced lung cancer planning.

Structure	Optimization Objective	Priority	Dose Values	Volume Difference[e]
PTV[a] in lung	1x Minimum	130	≥ 69Gy	-
	1x Maximum[d]	120	≤ 72Gy	-
PTV minus lung	1x Minimum	130	≥ 66Gy	-
	1x Maximum	120	≤ 72Gy	-
ITV[b]	1x Minimum	110	≥ 66Gy	-
Spinal Cord (+3mm)	1x Maximum	130 (110)	≤ 45 (48)Gy	-
Optimization objectives moved by AIO				
Contralateral lung	5x Upper	130	1.5 – 5.5Gy	22 – 37%
Total Lung – PTV	5x Upper	130	1.0 – 5.0Gy	22 – 37%
	2x Upper	80	18 – 19Gy	6%
Esophagus + 3mm PRV[c]	7x Upper	90	35 – 63Gy	13%
Heart	4x Upper	90	42 – 62Gy	12%
Skin	2x Upper	90	30 – 40Gy	6%

[a] PTV = Planning Target Volume

[b] ITV = Internal Target Volume

[c] PRV = Planning-at-Risk Volume

[d] Dose-volume objective at 0% volume

[e] Volume difference between the optimization objective and the respective DVH-line

Shoulder position variability between fractions may alter the delivered dose distributions as more or less dose is absorbed depending on their position. In an attempt to improve plan robustness against shoulder motion, direct irradiation through the shoulders (contoured as visible on the CT-scan, including a 2cm cranial expansion) was limited by including a maximum point dose objective (with an optimization priority of 110).

To avoid local under-dosage of PTVs near OARs, parts of the PTV_B and PTV_E that overlapped with a 5mm expansion of the included OARs were defined as extra PTV volumes to which separate minimum dose objectives were assigned. Table 3 summarizes

the disease characteristics and resulting PTV volumes of the included patients. A 5mm 'virtual' bolus region was routinely created where the PTVs extended outside the body[23]. A water-equivalent Hounsfield unit density value was assigned to this structure to allow for its inclusion during optimization. Dose reporting was done using the plans that incorporated this 'virtual' bolus region to ensure comparable PTV dose coverage and homogeneity values between all patients, regardless of the PTV volume extending outside the skin. To enforce dose fall-off outside PTV_B and PTV_E, 10mm ring structures were created, typically placed at 10mm / 5mm from PTV_B / PTV_E. When there was overlap, the PTV_E region, including a 5mm expansion, was subtracted from the ring around PTV_B. Additionally, the 'normal tissue objective' function was used with an optimization priority of 90, distance from the target border of 0.10cm, start and end dose values of 95% and 40%, respectively, and a dose fall-off value of 0.07.

OARs used during optimization could include the ipsilateral and contralateral parotid and submandibular glands; swallowing muscles[24] (upper esophageal sphincter, upper and lower larynx, superior, medial and inferior pharyngeal constrictor muscles, cricopharyngeal muscle and the trachea); and the oral cavity. If an OAR overlapped considerably with the PTV, the treating clinician could decide to sacrifice this OAR. Composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) structures, consisting of individually optimized OARs, were created solely for dose reporting purposes. For both parotid glands, the same optimization priorities and pulling distances were used since it would only become apparent throughout the interactive optimization process to what degree both OARs can be spared. This has been the clinical practice in our department for many years, based on the fact that for the majority of patients, ipsilateral parotid gland mean doses <40Gy can be achieved, which is a dose level for which still some functioning of this organ is maintained[25].

Manually interactively optimized plans (MPs)

Clinical coplanar RapidArc plans were created using either two full 6MV arcs (n=68) or two 270 degree 6MV arcs (n=2, chosen to reduce irradiation of the biopsy negative ipsilateral neck). The arcs rotated clockwise and counter-clockwise and typically used collimator angles of 40° and 45°. The clinical HNC plans were optimized by letting the planner manually adapt the optimization objectives interactively following our previously described protocol[22]. Serial OARs such as the brainstem, spinal cord and their planning at risk volumes (PRVs, incorporating a 3mm expansion) were assigned single fixed maximum dose objectives at 0% volume typically placed at 37Gy (39Gy for the PRVs), which is well below their respective dose tolerance levels. For the oral cavity, salivary glands and

swallowing muscles, 3-4 optimization objectives with the same weighting (priority), spread evenly along the dose-volume histogram (DVH)-line, were used to reduce the mean dose. Since the optimal locations of these dose-volume objectives are unknown at the start of planning for individual patients, the planner intermittently adapts their location trying to maintain a constant diagonal distance to the respective DVH-line that is displayed throughout the RapidArc optimization process. Because many OARs are included, manual interactive optimization of complex HNC patients is a labor-intensive process requiring the full attention of the planner. More than one optimization run is often required to produce the final clinical plan.

Table 3. Tumor site, stage and size of the planning target volumes (PTVs) of the included patients.

Disease Site	Patients	Stage Range	Mean ± StDev (Range)		
			PTV_B[a] (cm^3)	PTV_E[a] (cm^3)	PTV_T[a] (cm^3)
Oropharynx	40	T1N1 - T4aN2cM0	187.4 ± 90.6 (46.8 - 438.8)	352.3 ± 100.2 (67.2 - 608.0)	80.8 ± 55.6 (18.6 - 360.0)
(Supra-) Glottic Larynx	18	T1b - T4aN2c	158.3 ± 128.5 (34.1 - 536.1)	361.7 ± 90.0 (207.6 - 514.2)	54.7 ± 58.0 (10.1 - 258.4)
Hypopharynx	6	T2N0 - T3N2	256.2 ± 210.5 (39.1 - 607.0)	419.1 ± 116.8 (264.0 - 618.6)	83.0 ± 50.2 (26.1 - 138.9)
Nasopharynx	3	T1N2 - T1N2M0	217.6 ± 113.6 (120.4 - 342.5)	339.7 ± 61.5 (292.9 - 409.4)	106.3 ± 11.1 (93.8 - 114.8)
Unknown Primary	2	-	127.6 ± 42.0 (97.9 - 157.3)	445.3 ± 70.4 (395.5 - 495.1)	42.6 ± 7.6 (37.2 - 48.0)
Sinus Maxillaris	1	T4bN0M0	114.5	267.7	54.0

[a] PTV_B, PTV_E, PTV_T = Boost, elective and transition planning target volumes

AIO plans (APs)

The APs used the same field setup as their respective clinical plans. Ten optimization objectives, placed with equal volume spacing along the DVH-line, were used for the oral cavity, individual salivary glands and individual swallowing muscles (Table 1). All optimization objectives were adjusted by AIO approximately every 20 seconds and placed at a fixed diagonal distance below their respective DVH-line. Since reducing mean OAR doses is generally considered important for HNC patients, AIO kept the volume-values of the optimization objectives constant, whilst continuously striving the reduce the associated dose values. If the mean dose of an OAR was higher than 50Gy after the first multiresolution level of the optimizer, no further attempts to spare this OAR were made by automatically removing its dose-volume objectives from the optimization.

Study endpoints

The HNC plans were compared using the following dosimetric parameters: (1) Mean doses to individual OARs and the oral cavity, $comp_{sal}$ and $comp_{swal}$ (D_{oc}, D_{sal} and D_{swal}, respectively); (2) Maximum spinal cord and brainstem doses; (3) PTV volumes receiving 95% / 107% of the prescribed dose (V95% / V107%) and dose inhomogeneity (IH), defined for PTV_B / PTV_E as IH_B / IH_E = 100% − V95% + V107%; (4) The mean dose, and volume of the body − PTV structure receiving 5Gy, 30Gy and 50Gy (V5Gy, V30Gy and V50Gy, respectively); (5) The number of monitor units (MUs) required. Paired two-sided Student t-tests were performed with $p \leq 0.05$ considered significant.

One head and neck radiation oncologist performed a blinded evaluation of 20 APs and their corresponding MPs. These patients were considered to be a representative sample of the total of 70 patients because the set consisted of patients with the primary tumor located in the oropharynx (n=12), (supra-) glottic larynx (n=4), hypopharynx (n=2), nasopharynx (n=1), and unknown (n=1), showing a similar distribution as the 70 patient cohort (Table 3). Furthermore, the individual PTV volumes of these 20 patients were not found to be significantly different from the remaining set of 50.

Plan evaluation took into consideration institutional guidelines for PTV_B / PTV_E (aiming for V95% ≥ 99% / 98%). Regions of under-dosage would have to be small and located near the edges of the PTV to avoid 'cold spots' in the GTV. In general, no more than 4% / 15% of PTV_B / PTV_E would have to receive >107% of the prescription dose. Hot spots were avoided in the mandible or laryngeal cartilage. The remainder of the body was evaluated for the presence of 'hot spots' and dose fall-off outside the PTV.

Lung cancer plans

Previously generated[26], manually optimized RapidArc plans for 20 LC patients were selected for this analysis. The average PTV volume for this group was 927±337cm^3. The MPs and APs were created using two coplanar 6MV arcs, typically using 40 / 45 degree collimator angles, with an avoidance sector to avoid direct irradiation through the contralateral lung. Planning of LC patients was challenging because they included sparing of not only the lungs and spinal cord, but the heart and esophagus were also spared from medium to high doses. The skin, defined as the outer 5mm of the body contour (delineated only on the CT-slices containing the PTV) was also included during optimization to reduce the chance of skin toxicity[27]. 66Gy was prescribed to the PTV in 33 fractions of 2.0Gy.

Lung parameters considered important were the contralateral lung (CL) and total lung – PTV (TL, consisting of the contralateral and ipsilateral lungs with the PTV deducted) V5Gy and TL V20Gy[28]. To reduce these parameters, all optimization objectives were placed near the clinically relevant dose values for each OAR (Table 2). Throughout the optimization process, AIO maintained the initial dose-values of the optimization objectives, whilst attempting to reduce the associated volume values. A fixed vertical distance between the objectives and their respective DVH-line was therefore maintained. Contrary to AIO for HNC, where a constant 'pulling distance' (dose difference between the DVH-line and objective) was used for all OAR objectives, the 'pulling distance' for LC varied between the objectives in a way that was illustrative of the clinical choices made in our department regarding the different optimization goals. For example, the volume difference between the contralateral lung objectives (placed at 1.5-5.5Gy) and the DVH-line was 37-22%, while a volume difference of 12% was used for the heart objectives (placed at 42-62Gy). This was consistent for all patients. During the first optimization (i.e., not during the CPO), the PTV was divided into a PTV located in lung (low density) and the rest of the PTV (outside lung). Higher minimum dose objectives are assigned to the low density PTV structures during optimization to ensure that sufficiently high doses remain in this PTV after dose calculation. This separation of the PTV structure is not necessary during the CPO calculation because the AAA calculated dose is taken into account.

A comparison between the AP and MP LC plans was made considering: (1) PTV dose coverage (V95%) and homogeneity (using V107% and the homogeneity index [HI = 100% × (D2% − D98%) / D50%]). (2) Dose conformity (CI), defined as the PTV volume receiving 95% of the prescribed dose (PD) divided by the isodose volume receiving 95% PD. (3) CL V5Gy, and TL V20Gy, V5Gy and mean dose. (4) Heart and esophagus V60Gy, V40Gy and mean dose. (5) Skin volumes receiving 50Gy and 30Gy and body – PTV volumes receiving 60Gy and 70.6Gy (107% PD) (6) Number of monitor units. Paired two-sided Student t-tests were performed with $p \leq 0.05$ considered significant.

Results

Head and neck cancer plans

Results of the MPs and AIO plans for HNC are averaged over all 70 patients and summarized in Table 4. Maximum doses to the actual spinal cord and brainstem are reported, not to their planning-at-risk (PRV) volumes. Since only OARs that were intended to be spared in the original clinical plan were considered for this evaluation, the number of included OARs varied between the plans. 28 AIO plans (40%) required a CPO with higher optimization

priorities on the PTVs to improve PTV dose coverage and homogeneity values, ranging between 140 and 200 (150 was used in 19 out of 28 cases). These higher priorities during the CPO were solely chosen based on the discrepancy of the initial plan with the PTV dose coverage and homogeneity values as specified in our institutional planning protocols, i.e., the planner had no access to the original manually optimized plan, where a CPO could also have been run with higher priorities. In order to show the influence of the CPO with higher priorities, the dosimetric results of the plans before performing this additional CPO are indicated in Table 4 by 'AP_{1CPO}', while 'AP' represents the plans in which the additional CPO was performed, when necessary. The additional CPO improved PTV_B and PTV_E IH values by 0.7% and 1.8%, respectively, while increasing D_{sal}, D_{swal} and D_{OC} by 0.1-0.3Gy. These small differences show that the additional CPO was not a prerequisite for improved plan quality over the manually optimized plan using AIO. Given the small differences between the 'AP_{1CPO}' plans and 'APs', only the APs are considered for the further analysis.

Clinically acceptable maximum doses to the brainstem and spinal cord were achieved in all plans. Compared to the MPs, PTV_B dose inhomogeneity (IH_B) decreased significantly by 0.7% ($p<0.05$) in the APs, while IH_E decreased non-significantly. Except for the ipsilateral submandibular gland and the trachea, which were only included in 11 and 9 patients, respectively, mean dose to all individual parallel OARs decreased significantly in the APs over the MPs ($p<0.05$ to 0.0001). On average, mean dose to the individual salivary glands decreased by 0.8-1.2Gy, with D_{sal} decreasing by 1.0Gy. Swallowing muscles benefitted more using AIO, with average gains ranging from 2.7-6.0Gy, and D_{swal} decreasing by 4.0Gy. Finally, D_{oc} decreased by 1.6Gy on average in the APs. HNC APs required 2.9% more monitor units than the MPs average, suggesting that increased modulation was required to improve plan quality. For one patient, AIO resulted in a 5.1Gy increased ipsilateral parotid gland mean dose. This however was associated with mean dose reductions of 4.1Gy, 5.9Gy and 10.8Gy to the contralateral submandibular gland, oral cavity and $comp_{swal}$, respectively. Improved OAR sparing in the APs was associated with a slight, but statistically significant increase in dose to the body – PTV volume: On average, V5Gy, V30Gy and V50Gy increased by 0.3-0.5% in the APs, while the mean dose increased by 0.3Gy.

Figure 1a shows the final position of the optimization objectives of all included OARs in the MP (circles) and AP (crosses) for one patient in which AIO substantially improved OAR sparing. In general, the AP optimization objectives were placed at lower dose values compared to the MPs for all OARs, resulting in dose reductions to corresponding OARs. Additionally, more optimization objectives were used, resulting in dose reductions along the entire volume axis. Resulting D_{sal}, D_{swal} and D_{oc} values were

29.0Gy / 24.5Gy, 25.6Gy / 15.3Gy and 22.7Gy / 17.2Gy, respectively, in the MPs / APs, while PTV dose homogeneity was similar. In contrast, Figure 1b shows the optimization objectives in which the AP and MP both achieved similar quality. The MP objectives are located near the AP objectives, resulting in similar OAR sparing.

In 19 out of 20 (blinded) cases evaluated by the radiation oncologist, the AP was preferred over the MP. In one patient, both plans were judged of similar quality, providing comparable OAR sparing and PTV dose homogeneity. APs were preferred for various reasons, including: (i) improved oral cavity, salivary gland and / or swallowing muscle sparing, while achieving similar PTV dose homogeneity as the MP (n=10); (ii) in addition to improving OAR sparing, it also provided more homogeneous doses to at least one PTV (n=7); (iii) substantially improved sparing of the salivary glands and swallowing muscles, although this came at the cost of a small dose increase to the oral cavity, judged clinically acceptable (n=2). Figure 2 shows D_{sal}, D_{swal}, and D_{oc} values for all evaluated patients. In the APs for these patients, PTV_B / PTV_E IH improved by on average 0.6% / 2.9% over the MPs, while oral cavity / $comp_{sal}$ / $comp_{swal}$ mean doses improved significantly by 1.9Gy / 0.8Gy / 3.7Gy ($p<0.001$ to $p=0.03$). For the remaining patients (not included in the blinded comparison), similar differences were obtained, with PTV_B IH and oral cavity / $comp_{sal}$ / $comp_{swal}$ mean doses improving significantly by 0.7% and 1.4Gy / 1.1Gy / 4.0Gy ($p<0.001$ to 0.029), respectively, and PTV_E IH improving by 0.7%. Given the similarity of these improvements we therefore considered these 20 patients to be a representative sample for the entire cohort.

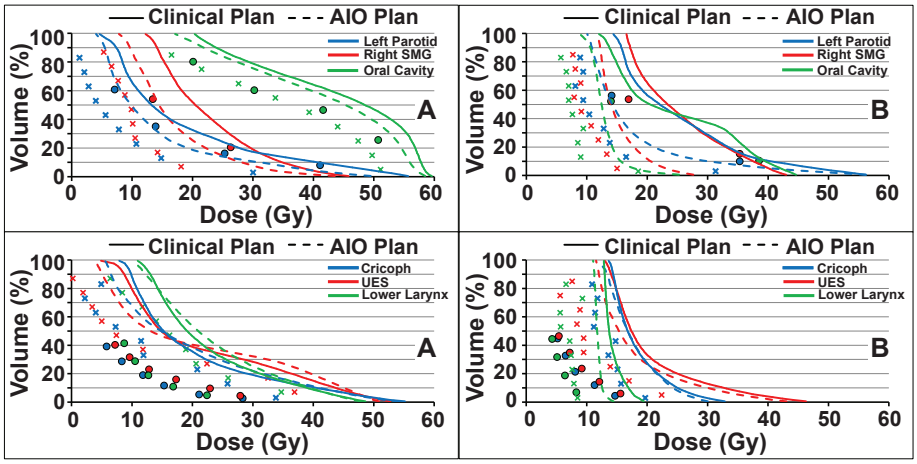


Figure 1. The final placement of the optimization objectives of all included organs-at-risk (OARs) for a patient in which the automatic interactive optimizer (AIO) plan substantially improved sparing over the clinical plan (Figure A), and for a patient in which similar plan quality was obtained (Figure B). Circles and crosses show the optimization objectives of the clinical and AIO plans, respectively.

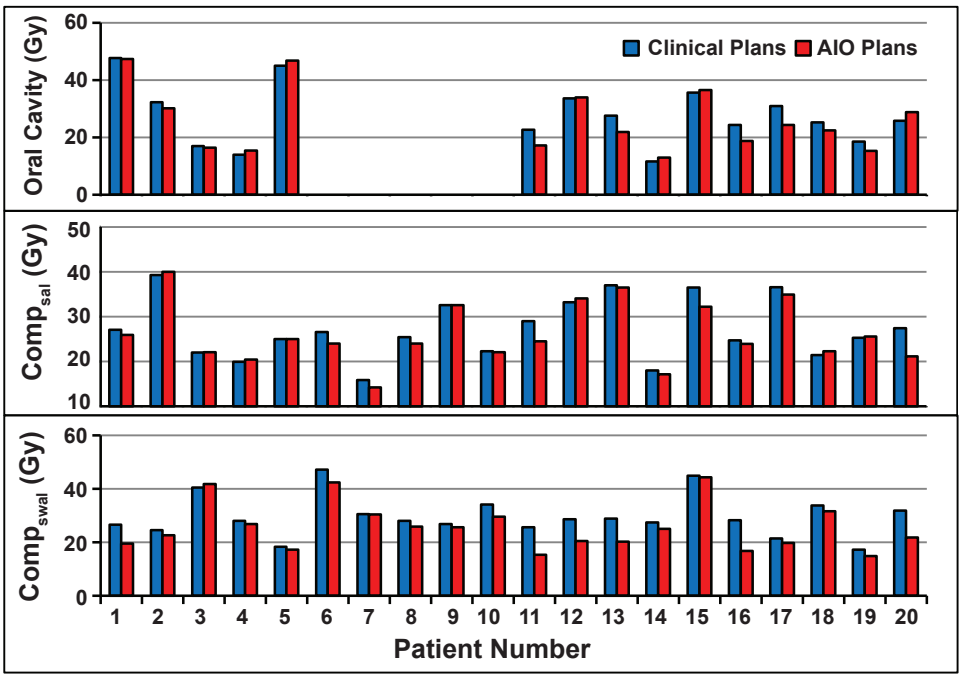


Figure 2. Mean dose to the oral cavity, salivary glands and swallowing muscles achieved in the clinical (blue) and automatic interactive optimizer (AIO, red) plans for the patients that were evaluated by the radiation oncologist.

Table 4. Dosimetric data for the manual (clinical) and automatic (AIO) interactively optimized treatment plans, averaged over all 70 patients.

Plan	Manual (MPs)	AP_{1CPO}	AIO (APs)	# Included
Boost planning target volume (%)				
V95%	99.1 ± 0.3	99.0 ± 0.8 (NS)[g]	99.3 ± 0.4 (p < 0.0001)	70
V107%	1.7 ± 2.9	1.8 ± 3.0 (NS)	1.3 ± 1.8 (NS)	70
IH[a]	2.7 ± 3.0	2.7 ± 3.5 (NS)	2.0 ± 1.9 (p < 0.05)	70
Elective planning target volume (%)				
V95%	98.0 ± 1.1	96.7 ± 5.6 (NS)	98.0 ± 0.6 (NS)	70
V107%	12.3 ± 7.2	11.5 ± 5.3 (NS)	11.0 ± 4.8 (NS)	70
IH[a]	14.3 ± 7.5	14.8 ± 8.2 (NS)	13.0 ± 4.8 (NS)	70
Max dose (Gy)				
Spinal Cord	39.0 ± 3.3	40.4 ± 2.6 (p < 0.01)	40.8 ± 1.4 (p < 0.0001)	70
Brainstem	34.6 ± 11.0	36.1 ± 8.0 (NS)	36.5 ± 8.6 (NS)	32
Mean dose (Gy)				
Contra. Parotid	19.0 ± 6.3	17.8 ± 6.1 (p <0 .0001)	18.1 ± 6.3 (p < 0.0001)	70
Ipsi. Parotid	26.3 ± 8.3	24.8 ± 7.9 (p < 0.0001)	25.1 ± 8.0 (p < 0.001)	66
Contra. SMG[b]	32.5 ± 7.9	31.4 ± 8.6 (p < 0.01)	31.7 ± 8.8 (p < 0.05)	55
Ipsi. SMG[b]	37.1 ± 7.7	35.9 ± 7.6 (NS)	36.2 ± 7.8 (NS)	11
Oral Cavity	25.4 ± 9.1	23.7 ± 8.9 (p < 0.001)	23.8 ± 8.8 (p < 0.001)	58
Cricoph[c]	25.6 ± 9.3	21.9 ± 9.0 (p < 0.0001)	22.3 ± 9.0 (p < 0.0001)	52
Lower Larynx	25.3 ± 10.8	20.8 ± 10.9 (p < 0.0001)	21.3 ± 11.1 (p < 0.0001)	44
Upper Larynx	32.4 ± 8.4	26.8 ± 8.0 (p < 0.0001)	27.3 ± 8.3 (p < 0.0001)	29
INF PCM[d]	30.3 ± 9.0	25.8 ± 9.5 (p < 0.0001)	26.4 ± 9.7 (p < 0.0001)	40
MED PCM[d]	44.5 ± 8.3	40.8 ± 9.0 (p < 0.0001)	41.3 ± 8.9 (p < 0.001)	26
SUP PCM[d]	40.9 ± 10.4	37.9 ± 11.4 (p < 0.0001)	38.2 ± 11.4 (p < 0.0001)	35
UES[e]	21.8 ± 9.2	17.9 ± 8.6 (p < 0.0001)	18.1 ± 8.7 (p < 0.0001)	62
Esophagus	24.4 ± 8.9	18.0 ± 6.5 (p < 0.01)	18.4 ± 7.1 (p < 0.01)	11
Trachea	26.5 ± 7.5	22.7 ± 4.6 (p < 0.05)	23.8 ± 5.2 (NS)	9
$Comp_{sal}$[f]	24.2 ± 6.5	22.9 ± 6.2 (p < 0.0001)	23.2 ± 6.3 (p < 0.0001)	70
$Comp_{swal}$[f]	29.5 ± 7.2	25.2 ± 7.1 (p < 0.0001)	25.5 ± 7.1 (p < 0.0001)	70
Body – PTV volume				
Mean dose (Gy)	10.8 ± 2.4	11.0 ± 2.4 (p < 0.0001)	11.1 ± 2.4 (p < 0.0001)	70
V5Gy (%)	41.8 ± 9.3	42.0 ± 9.3 (p < 0.05)	42.1 ± 9.4 (p < 0.05)	70
V30Gy (%)	14.3 ± 3.9	14.8 ± 4.0 (p < 0.001)	14.8 ± 4.0 (p < 0.001)	70
V50Gy (%)	2.5 ± 0.8	2.8 ± 0.9 (p < 0.0001)	2.9 ± 0.9 (p < 0.0001)	70
Monitor Units	509 ± 53	521 ± 37 (p < 0.05)	524 ± 37 (p < 0.001)	70

[a] IH = PTV dose inhomogeneity

[b] SMG= Submandibular gland

[c] Cricoph = Cricopharyngeal muscle

[d] PCM = Pharyngeal constrictor muscle

[e] UES = Upper esophageal sphincter

[f] $Comp_{sal}$ / $Comp_{swal}$ = composite salivary / swallowing structures

[g] NS = No significant differences with respect to manually optimized plan.

Lung cancer plans

Table 5 shows the dosimetric results, averaged over all 20 LC patients. The same OARs were included for sparing in all 20 LC patients. Similar to the HNC plans, an additional CPO with higher optimization priorities on the PTVs was necessary in 35% of 20 LC patients (n=7). In 4 and 3 cases respectively, PTV optimization priorities of 140 and 160 were chosen. The dosimetric results of the plans before performing this additional CPO are indicated in Table 5 by 'AP_{1CPO}', while 'AP' represents the plans in which the additional CPO was performed, when necessary. In line with the HNC plans, the additional CPO only marginally influenced the results, and only the LC APs are considered for further analysis.

Compared to the MPs, on average, the APs had significantly lower CL lung V5Gy (6.1%), heart V60Gy / mean dose (1.2% / 0.9Gy) and esophagus V60Gy / mean dose (2.8% / 0.9Gy). However, there were small, but statistically significant increases in maximum spinal cord dose (1.8Gy), Skin V50Gy / V30Gy (5.6cm^3 / 7.1cm^3) and Body – PTV V60Gy (56.5cm^3). Finally, TL V5Gy was 1.6% lower in the APs, while TL V20Gy was 1.4% higher, although these differences were not significant. Contrary to HNC, LC APs required 2.2% less monitor units compared to the MPs.

Although the mean improvements were small, larger trade-offs were apparent in some cases. In one patient, the AP improved the heart V60Gy, V40Gy and mean dose by 8.2%, 8.1% and 3.8%, respectively, at the cost of a 2.6% increase in TL V20Gy, and a near doubling of the Body – PTV volume receiving 60Gy. In another patient, TL V20Gy and mean dose in the AP were lower by 4.3% and 2.2Gy, respectively, while TL and CL V5Gy decreased by 6.8% and 9.5%. After review by the radiation oncologist, AIO was deemed sufficiently robust for clinical use in locally advanced lung cancer. This is supported by Table 5, which shows close similarity between the clinically relevant data of the MPs and APs.

Table 5. Dosimetric results for the Stage III lung cancer patients, averaged over all 20 evaluated patients.

Plan	MP	AP_{1CPO}	AP
Planning target volume			
V95% (%)	94.1 ± 1.2	93.4 ± 2.1	94.2 ± 1.2
V107% (%)	9.9 ± 6.5	8.8 ± 5.7	8.8 ± 5.4
D_{max} (Gy)	76.2 ± 1.6	75.7 ± 1.2	75.7 ± 1.1
D_{mean} (Gy)	67.4 ± 0.5	67.4 ± 0.4	67.4 ± 0.5
CI[a] (%)	0.73 ± 0.09	0.73 ± 0.07	0.72 ± 0.06
HI[b] (%)	17.5 ± 2.6	18.0 ± 0.04	17.3 ± 2.1
Lungs – PTV			
V20Gy (%)	24.4 ± 8.6	25.7 ± 7.3	25.8 ± 7.6
V5Gy (%)	54.2 ± 11.7	52.4 ± 10.1	52.6 ± 10.0
D_{mean} (Gy)	15.3 ± 3.6	15.8 ± 3.8	15.5 ± 4.1
CL Lung V5Gy (%)	39.1 ± 15.2	32.7 ± 12.4	33.0 ± 12.5*
Spinal cord D_{max} (Gy)	48.3 ± 2.6	50.0 ± 1.8*	50.1 ± 1.8*
Heart			
V60Gy (%)	8.9 ± 6.2	7.6 ± 5.8*	7.7 ± 5.8*
V40Gy (%)	21.6 ± 16.4	20.6 ± 15.0	20.6 ± 15.0
D_{mean} (Gy)	19.6 ± 11.2	18.6 ± 10.4*	18.7 ± 10.4*
Esophagus			
V60Gy (%)	23.7 ± 15.0	20.4 ± 14.2*	20.9 ± 14.5*
V40Gy (%)	41.7 ± 17.6	40.5 ± 17.8	40.7 ± 17.8
D_{mean} (Gy)	32.3 ± 9.3	31.3 ± 9.4*	31.4 ± 9.5*
Skin			
V50Gy (cm^3)	11.9 ± 12.3	16.2 ± 13.2*	17.5 ± 12.8*
V30Gy (cm^3)	112.6 ± 62.0	120.8 ± 71.0*	119.7 ± 72.6*
Body – PTV			
V60Gy (cm^3)	485.0 ± 260.1	536.8 ± 268.1*	541.5 ± 263.7*
V107% (cm^3)	14.7 ± 27.1	16.1 ± 18.6	14.7 ± 17.1
Monitor Units	721 ± 59	706 ± 57	705 ± 59

* Statistically significant difference ($p \leq 0.05$) with respect to the manual plan (two-sided Student t-test)

[a] CI = Dose conformity index

[b] HI = Planning target volume (PTV) homogeneity index

Discussion

This study presented detailed comparisons between manually optimized treatment plans and plans optimized using our in-house developed automatic interactive optimizer (AIO) for two sites that required a distinct approach to plan optimization.

AIO performance and optimal AIO settings were previously investigated on an initial group of 10 HNC patients[21]. Using the AIO settings from this initial work on a large group of 70 patients, we found that, on average, AIO improved both PTV dose homogeneity and OAR sparing over their respective MPs, showing that the AIO solution does not merely result in different trade-offs between the two[22]. These improvements were associated with slight increases in dose deposition to the remainder of the body.

Although the primary aim for HNC patients was reducing mean OAR doses, AIO has been designed to work with multiple dose-volume objectives per OAR, rather than a single mean dose objective. This is partly because the aim of AIO was to automate the interactive optimization process as it was performed manually in our clinic, and partly because automatically adjusting mean dose objectives will not result in good quality plans in treatment sites where a reduction of such doses is not considered clinically relevant (such as the locally advanced lung). AIO can therefore be applied to multiple treatment sites without necessitating a change in the automated interactive optimization process, apart from changing the angle at which the optimization objectives are placed with respect to the DVH-line. AIO was not designed to adjust the position of the maximum dose objectives, which remained at the same dose values throughout the entire optimization. This was done by design because in our clinic, it is generally considered more important to minimize mean doses values of the parallel OARs, rather than lowering the point dose values to the serials OARs as much as possible (as long as the maximum dose values to the serial OARs are well below their respective dose tolerance levels).

Although for HNC, average gains could be relatively small for individual OARs, the highly statistically significant values show that APs consistently improved OAR sparing over the MPs. The standard set of optimization objectives and priorities provides a fast and reliable approach to automatic optimization using AIO, although a subsequent CPO with higher optimization priorities on the PTVs could be required. Since for the majority of patients (60%) this additional CPO was not necessary, we believe that the set of optimization objectives used currently provides a reasonable starting point for the AIO optimization. As investigated previously[22], increasing the optimization priorities of the PTV objectives would come at the expense of OAR sparing. When using AIO in clinical practice, we

typically perform a CPO calculation with optimization priorities of 200 on the PTVs after creation of the initial plan. Although this approach slightly decreases the achievable levels of OAR sparing (Table 4), it does not require intermittent evaluation of the plan by the user and therefore increases the planning efficiency of the AIO solution.

For the 70 clinical HNC patients in the present study, an average of 2.4±1.2 (range: 1 to 9) plans were made to obtain the final plan. In this instance, each full new optimization run using different optimization objectives was considered to be a new plan (i.e., CPOs were not counted as such). In contrast, a full new optimization of the AP was not necessary for any patient. In addition, once AIO is started, user attendance is no longer required, allowing time to be invested in other activities. AIO therefore contributes to greater efficiency in treatment planning. All plans are subject to the same quality assurance process as MPs.

AIO was also successfully applied for the automatic optimization of locally advanced lung cancer, illustrating its versatility. For this site, we did not focus on reducing mean OAR doses, but rather attempted to reduce the OAR volumes receiving specific dose values. This required a different approach to automatic optimization compared to HNC. Contralateral lung V5Gy was noticeably lower in the LC APs (6.1%, on average), with more sparing of the heart and esophagus. Although these gains often came at the expense of statistically significant increases in dose deposition to the skin and remainder of the body, it should be noted that the absolute dose increases were typically small and likely not clinically relevant. It is important to note that MPs used for benchmarking the LC APs were all made by an experienced planning physicist, ensuring that AIO was given an appropriately challenging test. This is in contrast to the HNC MPs which were made by multiple treatment planners and thereby subject to more variation. For this reason, the a priori expectation was that differences between MPs and APs were likely to be more noticeable for HNC. Improved OAR sparing in the APs was likely achieved because; (i) AIO allows for more optimization objectives per OAR; (ii) the optimization objectives were adjusted more frequently and consistently than possible through manual interaction, allowing the optimization objectives to reach lower dose values (Figure 1). These findings can be generalized to manual interactive optimization using the Eclipse TPS.

Improved AP quality came at the expense of increased computing time for a single optimization. Depending on the number of OARs and size of the PTVs, computing times typically ranged from about 15-35 minutes compared to about 10-15 minutes for a single manual optimization without pausing or repeating optimization levels. In contrast to manual optimization, however, user attendance is not required during an AIO optimization

and planner time can be invested in other activities. Our experience so far is that for a standard HNC plan, a familiar user can usually complete the necessary AIO-specific preparatory work prior to optimization within 5 minutes, with the preparatory work including 1) selecting a template of optimization objectives, 2) indicating the boundaries of the optimization window to AIO, and 3) making AIO search the initial position of the optimization objectives in this window. This does not include the general preparatory work for planning (e.g. generating help structures, setting up the fields), the optimization time, or the time required to run a CPO if this is needed. This is comparable to 3.8±1.1 minutes of required working time presented in a recent study concerning Pinnacle Auto-Planning for VMAT HNC[29]. Breedveld et al.[7] noted that, depending on the number of IMRT beams, generating coplanar plans (including optimization) could take 1-2 hours. During this time however, optimal beam arrangements are also determined, while these need to be set up manually prior to an AIO optimization.

A number of approaches to automatic planning have been described with promising results. Knowledge-based planning requires the creation of a model with a sufficient number of plans that contain similar OARs, but varying OAR-PTV geometries[4,8,15,16,20,30–32]. This model can be used to predict the achievable DVHs for OARs and PTVs of new patients and use these to derive optimal optimization objectives. Some initial time and effort needs to be spent to ensure that high-quality plans are included in the model. Another automated planning solution is multi-criteria optimization (MCO), which typically uses a 'wish list' that contains PTV and OAR optimization priorities and weights[7,10,12–14]. Lexicographic optimization is similar to MCO and automatically adds optimization objectives to the optimization process[11]. Both approaches have been investigated and implemented for IMRT. As noted by Chen et al.[33] however, because of the rather complex computations, application to VMAT treatment planning is limited and reported studies either used relatively straightforward patient geometries[17,34] or used DVH information from MCO IMRT treatments to generate VMAT plans[33,35].

Potential limitations of this study include the following. When a large number of objectives (typically > 100) is used, the optimization screen may get cluttered, preventing optimal placement. This could be overcome by fully integrating AIO in the TPS. Although the AIO solution is specific to the Eclipse TPS, its core concept, gradually lowering OAR doses throughout the optimization process while the MLC apertures are being defined, may also be transferable to different TPSs. AIO merits further investigation for other types of tumors and clinical scenarios, as well as comparison with other automated planning methods. Finally, AIO specifically focuses on the automation of interactive optimization. Automating the entire planning process requires further work.

In conclusion, an in-house developed program that automates interactive optimization in the Eclipse TPS produced plans that were at least as good as, and often better, than manually optimized head and neck and lung cancer plans with a range of tumor sizes, locations and PTV-OAR geometry. When evaluated by a head and neck radiation oncologist, the automated HNC plans were preferred over their clinical counterpart in 95% of the cases. These results supported the implementation of AIO which has now been in routine clinical use in our department for RapidArc HNC planning since February 2014. AIO remains under evaluation for other sites. We also expect AIO to be useful for planning comparison studies because it removes potential user-dependent sources of bias[2].

References

1 J.I. Nord and J.Y. Peltola, *Interactive treatment plan optimization for radiation therapy. US Patent 7817778 B2*, (2010).

2 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

3 S. Everitt, T. Kron, N. Fimmell, J. Reynolds, C. Laferlita, D. Ball, M. Schneider-Kolsky, R. Budd, and M. Macmanus, "Interplanner variability in carrying out three-dimensional conformal radiation therapy for non-small-cell lung cancer.," J. Med. Imaging Radiat. Oncol. **52**(3), 293–6 (2008).

4 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt, "Patient geometry-driven information retrieval for IMRT treatment plan quality control.," Med. Phys. **36**(12), 5497–505 (2009).

5 Y. Matsuo, K. Takayama, Y. Nagata, E. Kunieda, K. Tateoka, N. Ishizuka, T. Mizowaki, Y. Norihisa, M. Sakamoto, Y. Narita, S. Ishikura, and M. Hiraoka, "Interinstitutional variations in planning for stereotactic body radiation therapy for lung cancer.," Int. J. Radiat. Oncol. Biol. Phys. **68**(2), 416–25 (2007).

6 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

7 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

8 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

9 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **79**(4), 1241–7 (2011).

10 D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning.," Med. Phys. **33**(9), 3399–407 (2006).

11 K.W. Jee, D.L. McShan, and B.A. Fraass, "Lexicographic ordering: intuitive multicriteria optimization for IMRT.," Phys. Med. Biol. **52**(7), 1845–61 (2007).

12 T.S. Hong, D.L. Craft, F. Carlsson, and T.R. Bortfeld, "Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head.," Int. J. Radiat. Oncol. Biol. Phys. **72**(4), 1208–14 (2008).

13 P.W.J. Voet, S. Breedveld, M.L.P. Dirkx, P.C. Levendag, and B.J.M. Heijmen, "Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT.," Med. Phys. **39**(8), 4858–65 (2012).

14 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

15 X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

16 B. Wu, T. McNutt, M. Zahurak, P. Simari, D. Pang, R. Taylor, and G. Sanguineti, "Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **84**(5), e647–53 (2012).

17 C. Boylan and C. Rowbottom, "A bias-free, automated planning tool for technique comparison in radiotherapy - application to nasopharyngeal carcinoma treatments.," J. Appl. Clin. Med. Phys. **15**(1), 4530 (2014).

18 E.M. Quan, J.Y. Chang, Z. Liao, T. Xia, Z. Yuan, H. Liu, X. Li, C.A. Wages, R. Mohan, and X. Zhang, "Automated volumetric modulated Arc therapy treatment planning for stage III lung cancer: how does it compare with intensity-modulated radio therapy?," Int. J. Radiat. Oncol. Biol. Phys. **84**(1), e69–76 (2012).

19 A.W.M. Sharfo, P.W.J. Voet, S. Breedveld, J.W.M. Mens, M.S. Hoogeman, and B.J.M. Heijmen, "Comparison of VMAT and IMRT strategies for cervical cancer patients using automated planning," Radiother. Oncol. (2015).

20 J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Evaluation of a knowledge-based planning solution for head and neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **91**(3), 612–20 (2015).

21 J.P. Tol, M. Dahele, J. Peltola, J. Nord, B.J. Slotman, and W.F.A.R. Verbakel, "Automatic interactive optimization for volumetric modulated arc therapy planning," Radiat. Oncol. **10**(1), 75 (2015).

22 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

23 C. Thilmann, A. Zabel, S. Nill, B. Rhein, A. Hoess, P. Haering, S. Milke-Zabel, W. Harms, W. Schlegel, M. Wannenmacher, and J. Debus, "Intensity-modulated radiotherapy of the female breast.," Med. Dosim. **27**(2), 79–90 (2002).

24 M.E.M.C. Christianen, J.A. Langendijk, H.E. Westerlaan, T.A. van de Water, and H.P. Bijl, "Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.," Radiother. Oncol. **101**(3), 394–402 (2011).

25 T. Dijkema, C.P.J. Raaijmakers, R.K. ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

26 G.J. Blom, W.F.A.R. Verbakel, M. Dahele, D. Hoffmans, B.J. Slotman, and S. Senan, "Improving radiotherapy planning for large volume lung cancer: a dosimetric comparison between hybrid-IMRT and RapidArc.," Acta Oncol. **54**(3), 427–32 (2015).

27 V.Y. Yazbeck, L. Villaruz, M. Haley, and M.A. Socinski, "Management of normal tissue toxicity associated with chemoradiation (primary skin, esophagus, and lung).," Cancer J. **19**(3), 231–7 (2013).

28 W.F.A.R. Verbakel, E. van Reij, I. Ladenius-Lischer, J.P. Cuijpers, B.J. Slotman, and S. Senan, "Clinical application of a novel hybrid intensity-modulated radiotherapy technique for stage III lung cancer and dosimetric comparison with four other techniques.," Int. J. Radiat. Oncol. Biol. Phys. **83**(2), e297–303 (2012).

29 J. Krayenbuehl, I. Norton, G. Studer, and M. Guckenberger, "Evaluation of an automated knowledge based treatment planning system for head and neck.," Radiat. Oncol. **10**(1), 226 (2015).

30 Y. Yang, E.C. Ford, B. Wu, M. Pinkawa, B. van Triest, P. Campbell, D.Y. Song, and T.R. McNutt, "An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection.," Med. Phys. **40**(1), 011709 (2013).

31 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

32 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

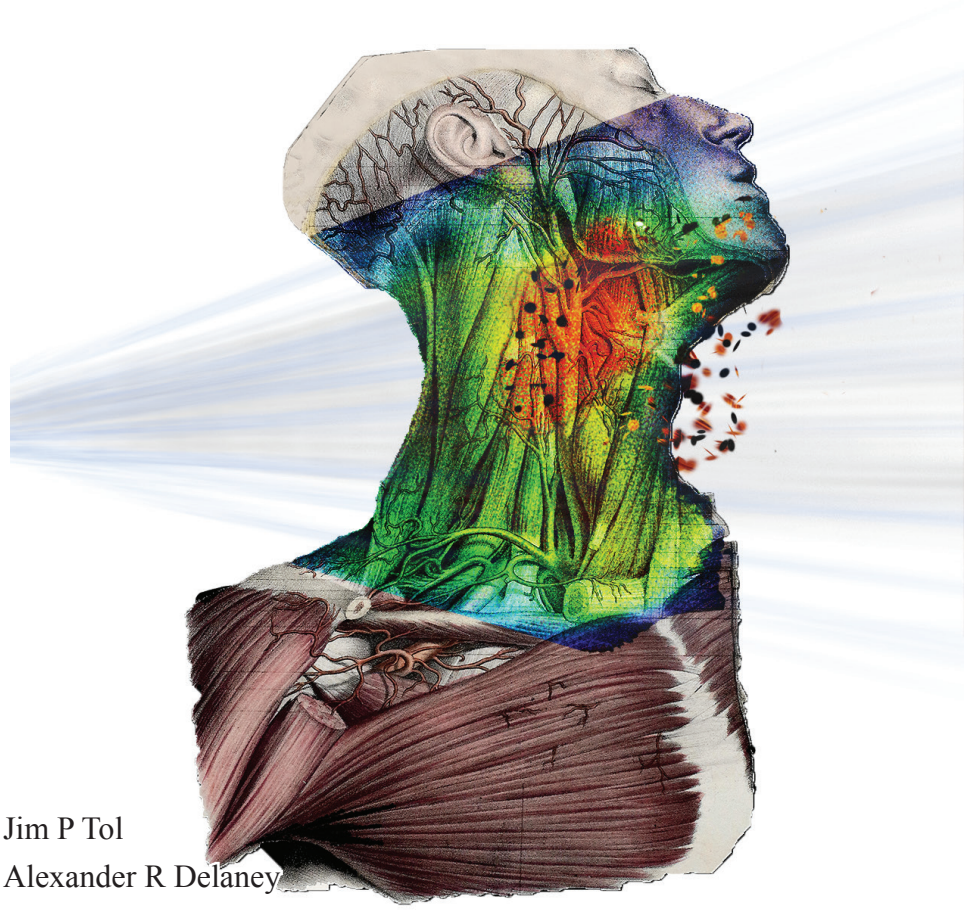
33 H. Chen, D.L. Craft, and D.P. Gierga, "Multicriteria optimization informed VMAT planning.," Med. Dosim. **39**(1), 64–73 (2014).

34 D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld, "Multicriteria VMAT optimization.," Med. Phys. **39**(2), 686–96 (2012).

35 B. Wu, D. Pang, P. Simari, R. Taylor, G. Sanguineti, and T. McNutt, "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study.," Med. Phys. **40**(2), 021714 (2013).

Chapter 7

Evaluation of a knowledge-based planning solution for head and neck cancer



Jim P Tol
Alexander R Delaney
Max Dahele
Ben J Slotman
Wilko FAR Verbakel

International Journal of Radiation Oncology, Biology, Physics **91**(3), 612-620 (2015).

Abstract

Purpose

Automated and knowledge-based planning techniques aim to reduce variations in plan quality. RapidPlan™ uses a library consisting of different patient plans to make a model that can predict achievable dose-volume histograms (DVHs) for new patients and uses those to place optimization objectives. We benchmarked RapidPlan versus clinical plans for two patient groups, using three different libraries.

Materials and Methods

Volumetric modulated arc therapy plans of 60 recent head and neck cancer patients that included sparing of the salivary glands, swallowing muscles and oral cavity, were evenly divided between two models, $Model_{30A}$ and $Model_{30B}$ and combined in a third model, $Model_{60}$. Knowledge-based plans were created for two evaluation groups; EG1, consisting of 15 recent patients, and EG2, consisting of 15 older patients in which only the salivary glands were spared. RapidPlan results were compared against clinical plans (CP) on boost / elective planning target volume homogeneity index (HI_B / HI_E = 100 × (D2% − D98%)/D50%) and mean dose to composite salivary glands, swallowing muscles and oral cavity (D_{sal}, D_{swal} and D_{oc}, respectively).

Results

For EG1, RapidPlan improved HI_B / HI_E values over CP by 1.0-1.3% / 1.0-0.6%. Comparable D_{sal} and D_{swal} values were seen in $Model_{30A}$ / $Model_{30B}$ / $Model_{60}$, decreasing by on average 0.1Gy / 1.0Gy / 0.8Gy and 4.8Gy / 3.7Gy / 4.4Gy, respectively. However, differences were noted between individual OARs, with $Model_{30B}$ increasing D_{OC} by 0.1Gy / 3.2Gy / 2.8Gy over CP / $Model_{30A}$ / $Model_{60}$. Plan quality was less consistent when the patient was flagged as an 'outlier'. For EG2, RapidPlan decreased D_{sal} by 4.1-4.9Gy on average, while HI_B / HI_E decreased by 1.1-1.5% / 2.3-1.9%.

Conclusions

RapidPlan knowledge-based treatment plans were comparable to CP if the patient's OAR-PTV geometry was within the range of those included in the models. EG2 results showed that a model including swallowing muscle and oral cavity sparing can be applied to patients with only salivary gland sparing. This may allow model library sharing between institutes. Optimal detection of inadequate plans and population of model libraries requires further investigation.

Introduction

Variation in knowledge and experience can lead to large differences in the quality of radiotherapy treatment plans[1,2] and may compromise the gains that can be realized with advanced technologies such as intensity modulated radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT). The same holds true for manpower and computing resources, which can affect the implementation of new treatment planning techniques and treatment planning capacity. Various solutions are being investigated to improve planning consistency[3–7], including increased automation of planning using knowledge-based approaches[8–13]. These typically use libraries of existing patient plans to create models that predict the amount of organ-at-risk (OAR) sparing that can be achieved for a new patient, based for example on the planning target volume (PTV)-OAR distance and overlap[14]. Being able to rationally predict OAR dose-volume histograms (DVHs) could remove the need for performing interactive optimization or multiple iterative optimizations, and could increase consistency in treatment planning[14–18]. Resulting plans produced by such a knowledge-based system should reflect the quality of the plans that populate the model and the ability of the software to predict the achievable DVHs.

RapidPlan™ (Varian Medical Systems, Palo Alto, USA) is a commercially available knowledge-based planning solution derived from previously published work[14,16]. Knowledge is in this case represented by models created from libraries of previous plans. The purpose of this report was: (i) to benchmark RapidPlan performance against recent clinical VMAT (RapidArc™, Varian Medical Systems, Palo Alto, USA) plans using model libraries made up of different plans and with different total numbers of plans, and (ii) to investigate whether model libraries based on plans that spare many OARs can be usefully applied to patients treated shortly after our clinical introduction of RapidArc, in whom fewer OARs were spared. This allowed us to test the versatility of the model libraries and to illustrate what a center new to RapidArc planning, and starting with the inclusion of only a few OARs, may anticipate when applying a model comprised of more advanced plans.

Materials and Methods

Clinical Plans

Clinical locally advanced head and neck cancer (HNC) treatment plans with a simultaneous integrated boost (SIB) were created using 6MV photons and 2 full arcs. Plans aimed to deliver 95% of the prescribed dose of 54.25-58.15Gy to 98% of the elective PTV (PTV_E; $V95 \geq 98\%$) and 95% of the prescribed dose of 70Gy to 99% of the boost PTV (PTV_B), in 35 fractions,

while limiting the volume of each PTV receiving >107% (V107) of the prescribed dose. A 5mm transition zone (PTV_T) was created between PTV_E and PTV_B to facilitate dose fall-off between them. OAR planning goals included maximum point doses to the spinal cord, brainstem and their planning at risk volumes (3mm expansion), and lowering the mean dose to parotid and submandibular glands (SMGs), individual swallowing muscles and the oral cavity as much as possible. HNC RapidArc optimization was performed interactively by continuously pulling the OAR optimization objectives according to our institutional optimization protocol[19–21]. Reduced clinical plan quality could occasionally be obtained if the institutional optimization protocol was not followed thoroughly.

Model Libraries

RapidPlan uses model libraries that contain dose distributions and OAR / PTV geometries of previously treated patients to generate a prediction range of achievable DVHs for individual OARs of new patients[22,23]. The optimization is automated by placing numerous dose-volume objectives along the lower range of the predicted DVHs. Although RapidPlan can calculate optimal priorities for optimization objectives, this feature is still being refined. A set of priorities reflecting our institutional practice was therefore entered manually.

Clinical RapidArc plans of 60 HNC patients treated between 2012-2014 (planned using the Eclipse treatment planning system v10.0.28) were arbitrarily selected and evenly divided among two models, $Model_{30A}$ and $Model_{30B}$, to investigate whether differences in the composition of plan libraries influence RapidPlan results. An additional model consisting of all 60 plans ($Model_{60}$) was used to evaluate the influence of model size. These models were used to create RapidArc plans for two evaluation groups using Eclipse v13.5, with the photon optimizer (PO) algorithm v13.5.10. These plans all used the same normal tissue objective settings. Dose calculation, with a subsequent 'continue previous optimization' (CPO) calculation to improve PTV dose homogeneity[19], was performed using the anisotropic analytical algorithm (AAA) v13.0.16 with a 2.5mm calculation grid. The knowledge-based plans were normalized to deliver the same mean dose to PTV_B as the respective clinical plan (CP).

To adequately predict DVHs of individual structures, RapidPlan requires each OAR to be present in at least 20 plans included in the model. However, some OARs in the individual plans were not actively spared because of the extent of OAR-PTV overlap. To avoid compromising a model by combining data of OARs that were not spared with DVHs of spared OARs, only those OARs that were actively spared in the CP were added to the model library. This meant that in some cases fewer than 20 individual OARs were available

from the 30 plans. As a result there was inadequate data to generate model estimations of the individual superior, medial and inferior pharyngeal constrictor muscles (PCMs) and the upper and lower larynx structures. To circumvent this, composite PCM and larynx structures were created from relevant individually spared OARs to estimate DVHs for individual PCM and larynx structures. These composite structures were both present in at least 20 plans in all models. Similarly, the contralateral SMGs were used to also model the ipsilateral SMGs.

Evaluation Groups

Evaluation group 1 (EG1) consisted of clinical treatment plans from 15 patients treated between 2013-2014, similar to those included in the models. EG1 was used to benchmark RapidPlan results against recent CPs. Optimization and dose calculation was performed using the progressive resolution optimizer (PRO) and AAA v10.0.28. OARs typically included the oral cavity, salivary glands and swallowing muscles. Depending on the degree of OAR-PTV overlap and whether or not the treating clinician chose to spare them, the salivary OARs could consist of some or all of the contra- and ipsilateral parotid glands and SMGs. Similarly, swallowing OARs could consist of some or all, of the upper esophageal sphincter, upper and lower larynx, superior, medial and inferior PCMs, cricopharyngeal muscle and esophagus.

Evaluation group 2 (EG2) consisted of plans from 15 patients patient treated in 2008-2009, planned using PRO / AAA v8.2.23-8.6.15. These plans were made at the beginning of our department's RapidArc program by a relatively inexperienced team. Attempts were made to spare the parotid and, less frequently, the submandibular glands, while the oral cavity and swallowing muscles were not spared.

The plans included in EG1 and EG2 were not included in the RapidPlan model libraries. To simplify OAR dose reporting, composite salivary and swallowing structures ($comp_{sal}$ and $comp_{swal}$, respectively), consisting of individually constrained OARs, were created. The size of the PTVs, $comp_{sal}$, $comp_{swal}$ and oral cavities included in the three models and two evaluation groups are shown in Table 1 while Table 2 shows the number of included OARs .

The included plans were made using different versions of the Eclipse treatment planning system, each using different optimization and dose calculation algorithms, and with minor differences in beam data used for the clinical and RapidPlan plans. To investigate whether differences between RapidPlan and CP results could originate from the use of different optimization algorithms, the CPs of 5 patients from 2008-2009 were recreated using PRO and AAA v10.0.28 and using PO v13.5.10 and AAA v13.0, without altering the optimization objectives.

Table 1: Size of the planning target volumes (PTVs), composite salivary / swallowing structures ($comp_{sal}$ / $comp_{swal}$) and oral cavities of the patients included in the three RapidPlan models ($Model_{30A}$, $Model_{30B}$ and $Model_{60}$) and two evaluation groups (EG1 and EG2).

Volume (cm^3)	PTV_B [a]	PTV_E [a]	PTV_T [a]	$comp_{sal}$ [b]	$comp_{swal}$ [c]	oral cavity
		$Model_{30A}$				
Mean	208.1 ± 119.6	346.9 ± 74.8	74.0 ± 34.0	66.2 ± 22.7	22.6 ± 15.1	70.1 ± 33.8
Range	39.1 to 607.0	223.5 to 514.2	17.6 to 155.7	20.7 to 118.2	4.8 to 67.5	14.7 to 186.6
		$Model_{30B}$				
Mean	150.4 ± 102.1	390.3 ± 102.4	60.8 ± 51.3	65.8 ± 19.4	25.4 ± 14.1	108.2 ± 68.0
Range	34.1 to 536.1	240.6 to 618.6	10.1 to 258.4	22.6 to 101.7	3.4 to 57.7	29.4 to 283.5
		$Model_{60}$				
Mean	179.3 ± 114.0	368.6 ± 91.6	67.7 ± 43.2	66.0 ± 21.0	24.0 ± 14.6	88.4 ± 55.9
Range	34.1 to 607.0	223.5 to 618.6	10.1 to 258.4	20.7 to 118.2	3.4 to 67.5	14.7 to 283.5
		EG1[d]				
Mean	236.7 ± 150.2	376.8 ± 108.3	89.0 ± 45.2	67.3 ± 27.5	27.0 ± 11.1	87.1 ± 55.3
Range	69.9 to 666.7	207.6 to 658.6	27.0 to 199.1	23.6 to 105.7	5.1 to 46.3	35.1 to 241.9
		EG2[e]				
Mean	200.6 ± 108.5	367.1 ± 121.7	51.6 ± 37.8	57.0 ± 18.8	-	-
Range	75.5 to 475.6	226.2 to 657.4	15.9 to 156.1	29.8 to 97.1	-	-

[a] Boost, elective and transition planning target volumes

[b] Composite salivary glands

[c] Composite swallowing muscles

[d] The first evaluation group, consisting of 15 patients treated at our institute between 2012 and 2014

[e] The second evaluation group, consisting of 15 patients treated at our institute in 2008 or 2009

Outlier Detection

The models should ideally consist of a variety of OAR / PTV geometries and variation in OAR dosimetry, resulting solely from geometric variations and not from large variations in plan quality. In RapidPlan's model configuration, the user can verify that the spread of OAR DVHs in the model results from the varying OAR / PTV geometries and not from poor or inconsistent planning quality. Principal component analysis is used to determine the geometric feature that correlates strongest with the spread of dosimetry in the model plans. This can allow the user to reject certain plans (e.g. those in which insufficient OAR sparing was achieved) to try and improve the model. This warning system may be further refined in future versions and it should be used in conjunction with visual and dosimetric plan evaluation when deciding whether to reject a plan from the model. Using this strategy, 6 / 60 patients in $Model_{60}$ ultimately identified as containing one outlier OAR. Nevertheless, these patients were included in the models to investigate RapidPlan performance providing a set of arbitrarily selected clinical plans, without removal / replanning of outlier plans.

Table 2: Number of organs-at-risk included in the three RapidPlan models ($Model_{30A}$, $Model_{30B}$ and $Model_{60}$) and two evaluation groups (EG1 and EG2).

	$Model_{30A}$	$Model_{30B}$	$Model_{60}$	EG1[e]	EG2[f]
Oral Cavity	27	24	51	15	-
CL parotid[a]	30	30	60	15	15
IL parotid[a]	27	29	56	15	15
CL SMG[b]	25	25	50	10	6
IL SMG[b]	5	5	10	2	4
Cricopharyngeal Muscle	24	23	47	11	-
Lower Larynx	19	20	39	12	-
Upper Larynx	14	14	28	9	-
Inferior PCM[c]	19	19	38	9	-
Medial PCM[c]	14	13	27	4	-
Superior PCM[c]	14	18	32	6	-
UES[d]	30	26	56	14	-

[a] Contralateral and ipsilateral parotid glands

[b] Contralateral and ipsilateral submandibular glands

[c] Pharyngeal constrictor muscles

[d] Upper esophageal sphincter

[e] The first evaluation group, consisting of 15 patients treated at our institute between 2012 and 2014

[f] The second evaluation group, consisting of 15 patients treated at our institute in 2008 or 2009

Before applying a model to calculate OAR DVH estimations for a patient, RapidPlan evaluates whether the OAR is an outlier according to a number of model parameters, or whether the OAR's geometry is located in a poorly populated region of the model. The evaluation parameters are OAR and PTV (combined as PTV_B + PTV_E + PTV_T) volumes, OAR-PTV distance and overlap, and the first geometric principal component score which correlates OAR-PTV geometry to OAR dosimetry[12]. An OAR is designated as an outlier when one or more of these metrics lie outside the range of values found in the model. This is intended to alert the user to the possibility of unreliable DVH estimations that might lead to sub-optimal plans.

Study Endpoints

A comparison between the knowledge-based plans and their respective CPs was made on the basis of (i) the homogeneity index (HI), calculated for PTV_B / PTV_E (HI_B / HI_E) using: HI = 100% × (D2% – D98%)/D50%, (ii) mean doses to individual OARs, $comp_{sal}$, $comp_{swal}$ and the oral cavity (D_{sal}, D_{swal} and D_{OC}, respectively). Paired, two sided Student t-tests were performed to identify significant differences ($p \leq 0.05$) between RapidPlan and clinical plans.

Results

In EG1, an average of 1.8 / 1.1 / 0.5 OARs per patient were flagged as outliers by $Model_{30A}$ / $Model_{30B}$ / $Model_{60}$, out of an average of 8.3 OARs per patient. In EG2 there were 0.4 / 0.5 / 0.3 outlier OARs per patient out of 2.7 OARs per patient. $Model_{60}$ resulted in the least outliers, indicating that this model can account for a larger range of patient geometries. Although we analyzed the outlier warnings given when applying the different models, investigating the influence of outliers on RapidPlan performance is beyond the scope of this study.

All knowledge-based plans were reviewed and deemed satisfactory by a senior clinical physicist experienced in HNC planning. Table 3 summarizes RapidPlan results for EG1, averaged over all patients, and Figure 1 shows results for individual patients. Compared with the CPs, PTV_B coverage (V95), V107 and HI_B / HI_E values improved using RapidPlan, although most differences were not significant. On average, all three models provided comparable D_{sal} and D_{swal}. Compared to the CP, D_{swal} decreased by 3.7-4.8Gy using RapidPlan. This was mostly caused by on average 4.3-6.0Gy more sparing of the upper and lower larynx and inferior PCM. Differences between the models were noted for some individual OARs. For example, $Model_{30B}$ resulted in an average D_{OC} of 26.5Gy; 0.1Gy / 3.2Gy / 2.8Gy higher than in CP / $Model_{30A}$ / $Model_{60}$.

The following outlier examples for EG1 patients are highlighted. (1) In patient #11 (see Figure 1) which had a large PTV_B volume (666.7cm^3), the combined target volume (PTV_B + PTV_E + PTV_T = 1137.9cm^3) and OAR-PTV overlap volume were flagged as outliers for all 8 OARs using $Model_{30A}$, while only the OAR-PTV overlap volume was flagged as an outlier for 4 OARs using $Model_{30B}$ and $Model_{60}$. D_{sal}, D_{swal} and D_{OC} results with $Model_{30A}$ / $Model_{30B}$ / $Model_{60}$ were 41.0Gy / 44.3Gy / 44.5Gy, 21.6Gy / 29.2Gy / 23.6Gy and 30.6Gy / 50.1Gy / 28.3Gy, respectively; (2) In patients #6 and #3, all included OARs were flagged as target volume outliers using $Model_{30A}$, while $Model_{30B}$ / $Model_{60}$ flagged 25% / 25% and 0% / 0% of OARs as outliers in these patients. In patient #6, highly variable plan quality was obtained, with maximum differences in D_{sal} / D_{swal} / D_{OC} of 1.8Gy / 8.0Gy / 11.3Gy between models. The worst performing model, $Model_{30B}$, increased D_{OC} by 8.1Gy over the CP. In contrast, D_{sal} / D_{swal} / D_{OC} varied in patient #3 by 3.1Gy / 1.3Gy / 4.4Gy between models and the gains over the CP were similar to those demonstrated by the pooled data.

A number of findings illustrate different trade-offs in the knowledge-based planning results. In patient #5, the models resulted in decreased D_{OC} (3.8-7.0Gy) and D_{sal} (3.6-4.6Gy), at the cost of higher HI_B (9.6-10.4% versus 8.4% clinically) and HI_E (12.8-13.4% versus 11.2%). In patient #10, $Model_{30A}$ improved D_{OC} by 7.8Gy compared to the clinical plan, at the cost of 2.3Gy higher D_{sal}. In contrast, $Model_{30B}$ / $Model_{60}$ decreased D_{OC} by 0.7Gy / 4.5Gy while D_{sal} reduced by 0.6Gy / 1.3Gy. Figure 2 shows DVH results and dose distributions of patient #1, in which the plan made by RapidPlan resulted in substantially improved $comp_{swal}$ sparing over the CP but also decreased dose conformity outside the PTVs. Visual inspection of the CP revealed deviation from the planning protocol with too few optimization objectives or sub-optimally located objectives for the individual swallowing muscles. The same was true for patient #14.

Results obtained for EG2, averaged over all 15 patients, are summarized in Table 4. Pooled data showed significant improvements in PTV_B coverage and sparing of both parotid glands using RapidPlan. Compared to the CPs, RapidPlan decreased HI_B / HI_E by 1.1-1.5% / 1.9-2.3%, while D_{sal} decreased (significantly) by 4.1-4.9Gy. In five patients, all 3 models achieved >6Gy reduction in D_{sal} compared to the CP.

The effect of different optimization and dose calculation algorithms on OAR and PTV dosimetry (averaged for 5 patients) is summarized in Table 5. Newer optimization algorithms primarily improved PTV dose homogeneity, while OAR sparing was less influenced.

After importing all plans and assigning the OAR and PTV structures, computing times to create $Model_{30A}$, $Model_{30B}$ and $Model_{60}$ were 19.1, 19.5 and 39.2 minutes, respectively. Generating DVH predictions took between 22 and 36 seconds per patient, depending on the number of OARs.

Table 3. RapidPlan results obtained over the first evaluation group (EG1), created using clinical plans of recently treated patients (2012-2014). Results are averaged over all 15 patients. The range shows the smallest and largest deviation when computing clinical minus RapidPlan results.

Plan	Clinical	$Model_{30A}$	$Model_{30A}$ (Range)	$Model_{30B}$	$Model_{30B}$ (Range)	$Model_{60}$	$Model_{60}$ (Range)
Boost PTV (%)							
V95	99.0 ± 0.0	99.3 ± 0.6	-1.0 to 1.1	99.5 ± 0.4*	-1.0 to 0.7	99.5 ± 0.5*	-1.0 to 0.8
V107	2.6 ± 3.9	2.0 ± 2.5	-6.7 to 7.9	1.4 ± 2.2	-1.4 to 7.8	1.5 ± 2.1	-1.8 to 7.7
HI[a]	10.3 ± 1.2	9.3 ± 1.1	-2.0 to 4.5	9.0 ± 1.3	-1.2 to 5.2	9.2 ± 1.3	-1.6 to 4.7
Elective PTV (%)							
V95	97.9 ± 0.8	97.8 ± 1.5	-2.3 to 2.9	98.2 ± 1.3	-2.7 to 2.7	98.1 ± 1.0	-2.2 to 1.7
V107	13.6 ± 9.0	11.2 ± 4.2	-10.3 to 22.6	9.7 ± 3.8	-4.6 to 24.6	9.9 ± 4.0	-4.0 to 23.4
HI[a]	15.4 ± 2.3	14.7 ± 1.7	-2.6 to 4.2	14.4 ± 1.7	-2.2 to 4.6	14.8 ± 1.6	-2.9 to 4.2
Max dose (Gy)							
Spinal Cord	40.5 ± 5.1	40.4 ± 0.7	-6.9 to 1.3	40.3 ± 0.8	-6.8 to 0.7	40.4 ± 0.9	-7.1 to 1.2
Brainstem	40.1 ± 8.0	38.2 ± 3.8	-6.5 to 14.0	38.8 ± 2.2	-2.8 to 13.9	37.9 ± 3.5	-7.0 to 13.2
Mean dose (Gy)							
Oral Cavity	26.4 ± 11.2	23.3 ± 8.9*	-2.3 to 8.3	26.5 ± 10.8	-11.2 to 9.3	23.7 ± 9.8*	-0.7 to 7.4
CL Parotid[b]	20.3 ± 4.3	20.0 ± 4.1	-3.1 to 2.5	18.4 ± 4.2*	-3.4 to 4.4	19.1 ± 3.6*	-1.2 to 3.9
IL Parotid[b]	27.0 ± 5.7	27.2 ± 6.7	-3.2 to 7.5	27.1 ± 6.9	-4.2 to 5.9	26.9 ± 7.1	-6.1 to 5.4
CL SMG[c]	33.2 ± 5.1	33.1 ± 4.5	-4.0 to 4.3	31.2 ± 6.8	-1.1 to 9.3	31.3 ± 6.3	-1.5 to 6.9
IL SMG[c]	52.3 ± 8.5	46.8 ± 9.3	4.9 to 6.1	47.7 ± 10.9	2.8 to 6.3	47.2 ± 10.8	3.5 to 6.8
Cricoph[d]	27.6 ± 10.5	26.1 ± 11.0	-7.0 to 15.8	25.8 ± 9.1	-8.9 to 15.9	25.5 ± 9.8	-4.5 to 18.5
Lower Larynx	27.5 ± 9.0	23.1 ± 9.3	-3.7 to 21.7	22.6 ± 8.7*	-1.3 to 20.0	22.5 ± 9.6*	-3.3 to 20.0
Upper Larynx	40.0 ± 12.5	34.5 ± 10.8*	0.6 to 12.2	34.7 ± 10.4*	-0.4 to 12.3	34.3 ± 10.8*	-0.3 to 11.3
Inferior PCM[e]	30.9 ± 8.2	25.2 ± 4.9*	-2.3 to 14.9	25.4 ± 4.3*	-2.0 to 15.3	25.1 ± 4.6*	-0.3 to 14.1
Medial PCM[e]	52.0 ± 8.1	50.6 ± 7.6	-1.7 to 3.6	51.4 ± 7.0	-2.2 to 3.3	51.6 ± 6.7	-3.9 to 3.5
Superior PCM[e]	44.8 ± 9.0	40.9 ± 8.7*	1.0 to 5.3	42.1 ± 9.7*	-0.1 to 5.6	40.4 ± 9.1*	1.4 to 6.2
UES[f]	23.3 ± 10.3	19.9 ± 7.3	-4.0 to 16.2	19.6 ± 10.5*	-4.5 to 18.3	19.8 ± 9.0*	-3.6 to 16.7
$Comp_{sal}$[g]	24.6 ± 4.3	24.5 ± 4.5	-2.3 to 3.6	23.6 ± 4.3	-2.5 to 4.6	23.8 ± 4.3	-2.5 to 4.6
$Comp_{swal}$[g]	32.9 ± 7.9	28.1 ± 5.9*	-2.7 to 14.2	29.2 ± 6.7*	-0.8 to 11.4	28.5 ± 6.3*	-1.4 to 11.8

* Statistically significant difference ($p \leq 0.05$) with clinical plan result

[a] boost and elective PTV homogeneity indices

[b] Contralateral and ipsilateral parotid glands

[c] Contralateral and ipsilateral submandibular glands

[d] Cricopharyngeal muscle

[e] Pharyngeal constrictor muscle

[f] Upper esophageal sphincter

[g] Composite salivary and swallowing structures

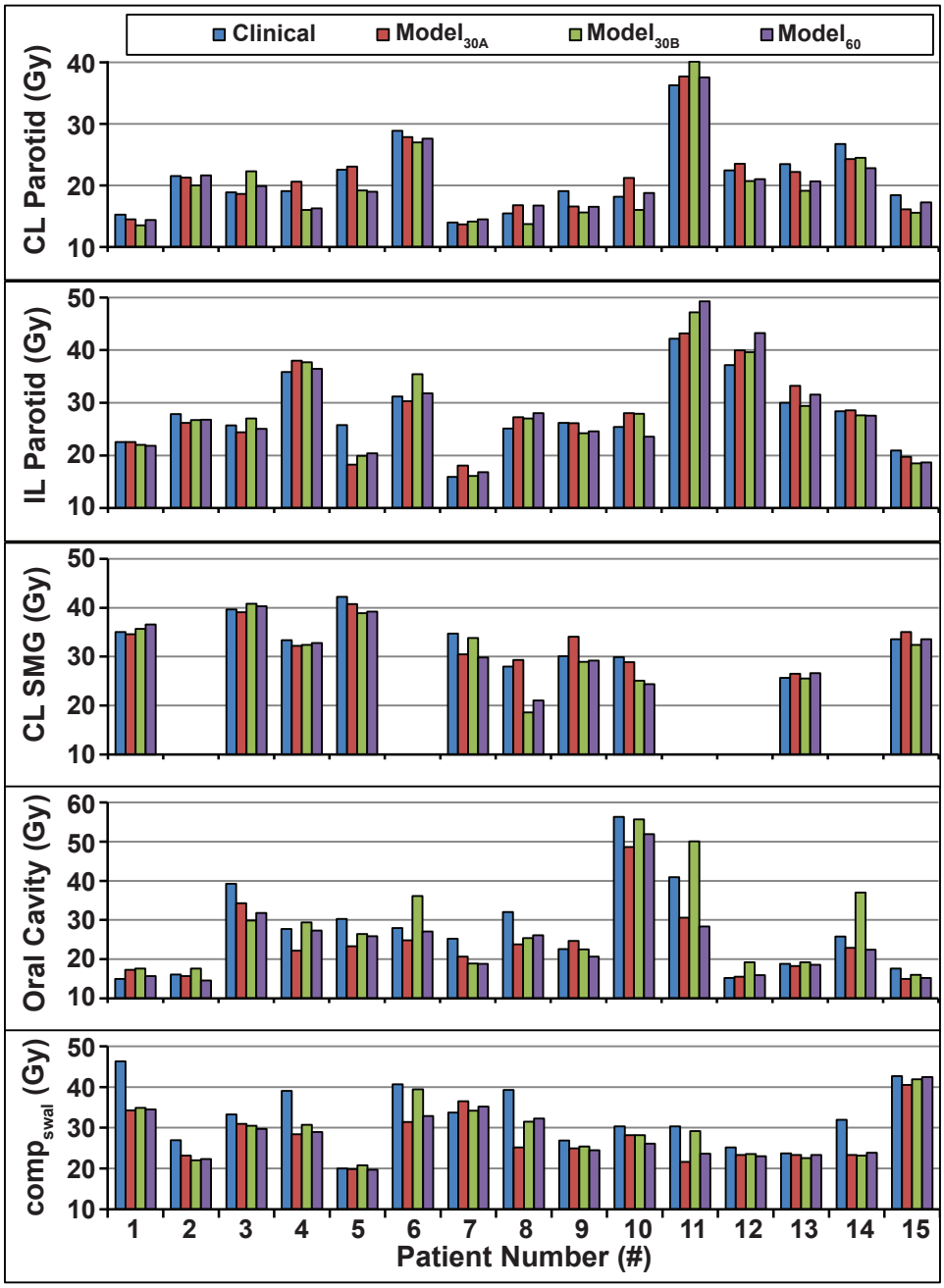


Figure 1. Histograms showing mean doses to the contralateral (CL) and ipsilateral (IL) parotid glands, contralateral submandibular gland (SMG), oral cavity and composite swallowing muscles ($comp_{swal}$) for the three knowledge-based plans and the clinical plan of all patients in evaluation group 1 (EG1).

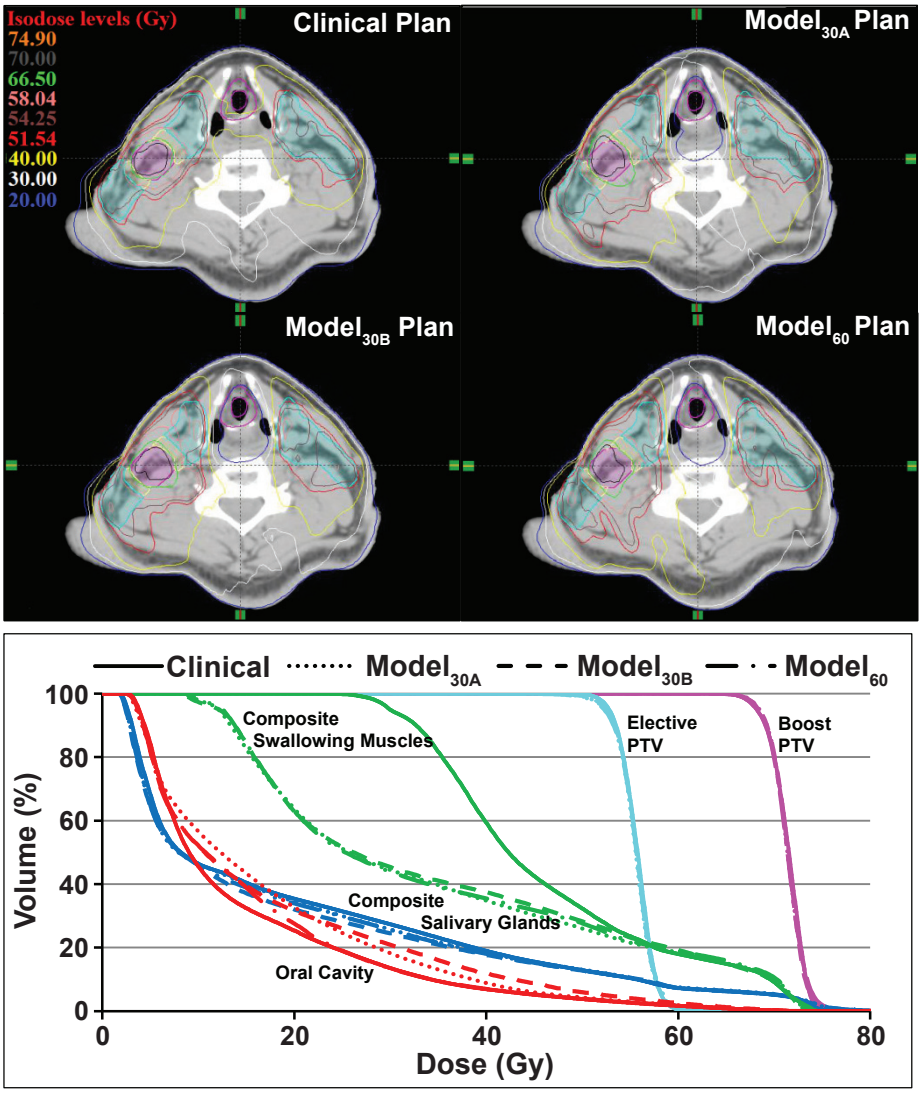


Figure 2. Dose distributions and dose-volume histograms (DVHs) of the clinical plan and three knowledge-based plans for a patient from EG1 in which improved sparing of the composite swallowing muscles (green DVH-lines) was obtained at the cost of lower dose conformity outside the boost and elective planning target volumes (PTVs, shaded magenta and cyan, respectively).

Table 4. Organ-at-risk (OAR) and planning target volume (PTV) dosimetry obtained when applying $Model_{30A}$, $Model_{30B}$ and $Model_{60}$ to the second evaluation group (EG2), consisting of patients treated between 2008-2009. Data is averaged over all 15 patients. The range shows the smallest and largest deviation when computing clinical minus RapidPlan results.

Plan	Clinical	$Model_{30A}$	$Model_{30A}$ (Range)	$Model_{30B}$	$Model_{30B}$ (Range)	$Model_{60}$	$Model_{60}$ (Range)
Boost PTV (%)							
V95	99.3 ± 0.5	99.8 ± 0.2*	-1.9 to 0.2	99.8 ± 0.2*	-2.0 to 0.3	99.7 ± 0.2*	-1.9 to 0.3
V107	1.7 ± 2.9	0.8 ± 1.5	-0.6 to 6.3	1.0 ± 1.6	-0.6 to 6.2	1.2 ± 2.2	-3.0 to 5.4
HI[a]	8.7 ± 1.7	7.2 ± 1.2*	-0.3 to 4.3	7.5 ± 1.4*	-1.7 to 3.9	7.6 ± 1.5	-1.3 to 3.5
Elective PTV (%)							
V95	97.8 ± 3.1	99.3 ± 0.7	-10.5 to 1.1	99.2 ± 0.9*	-8.8 to 0.7	99.3 ± 0.5	-10.7 to 0.7
V107	7.9 ± 9.6	4.5 ± 4.0	-3.5 to 22.9	5.9 ± 4.9	-5.7 to 14.7	5.8 ± 6.0	-5.8 to 13.6
HI[a]	13.0 ± 4.0	10.8 ± 2.3*	-1.3 to 8.9	11.1 ± 2.8	-1.5 to 6.2	10.7 ± 2.3*	-0.9 to 9.3
Max dose (Gy)							
Spinal Cord	43.4 ± 6.8	43.1 ± 6.2	-5.3 to 3.6	43.7 ± 6.7	-5.5 to 4.0	43.9 ± 6.7	-4.7 to 5.4
Brainstem	31.1 ± 16.5	30.7 ± 16.8	-2.2 to 4.6	30.2 ± 16.7	-1.2 to 5.5	30.9 ± 16.7	-2.5 to 5.1
Mean dose (Gy)							
CL Parotid[b]	23.9 ± 6.5	20.2 ± 6.5*	-0.6 to 8.9	19.1 ± 5.7*	1.3 to 10.7	19.5 ± 5.9*	-1.1 to 10.9
IL Parotid[b]	31.5 ± 8.8	27.4 ± 7.9*	-2.6 to 10.4	27.1 ± 8.9*	-2.0 to 12.2	26.5 ± 7.5*	-1.4 to 10.8
CL SMG[c]	40.6 ± 11.0	36.6 ± 12.4	-2.7 to 19.4	36.3 ± 12.2	-3.0 to 19.6	35.4 ± 12.6	-3.2 to 19.3
IL SMG[c]	50.1 ± 16.7	49.6 ± 19.0	-1.0 to 4.3	52.0 ± 15.8*	-0.9 to -2.8	47.0 ± 18.0	-1.0 to 8.6
$Comp_{sal}$[d]	28.8 ± 6.2	24.7 ± 5.7*	-1.3 to 8.2	24.2 ± 6.0*	0.2 to 9.7	23.9 ± 5.3*	-1.1 to 8.6

* Statistically significant difference ($p \leq 0.05$) with clinical plan result

[a] boost and elective PTV homogeneity indices

[b] Contralateral and ipsilateral parotid glands

[c] Contralateral and ipsilateral submandibular glands

[d] Composite salivary glands

Table 5. The influence using different versions of the progressive resolution optimizer (PRO), photon optimizer (PO) and anisotropic analytical algorithm (AAA) calculation models on planning target volume (PTV) dose homogeneity and organ-at-risk sparing.

Algorithm	PRO & AAA version 8.6-8.9	PRO & AAA version 10.0.28	PO & AAA version 13.0
Homogeneity indices (%)			
Boost PTV	9.5 ± 2.2	7.4 ± 1.5	7.5 ± 1.7
Elective PTV	13.1 ± 2.4	10.2 ± 2.3	10.6 ± 2.3
Mean dose (Gy)			
CL Parotid[a]	19.7 ± 1.4	19.3 ± 1.8	18.9 ± 2.0
IL parotid[a]	26.8 ± 5.0	26.2 ± 4.6	25.7 ± 4.5
CL SMG[b]	37.3 ± 6.8	36.9 ± 6.4	37.3 ± 7.7
IL SMG[b]	52.1 ± 12.2	51.4 ± 12.2	51.7 ± 12.4
$Comp_{sal}$[c]	27.5 ± 5.1	27.0 ± 5.5	26.7 ± 4.8

[a] Contralateral and ipsilateral parotid glands

[b] Contralateral and ipsilateral submandibular glands

[c] Composite salivary glands

Discussion

In this study we evaluated the performance of a commercial knowledge-based planning solution. HNC patients were chosen for the analysis because their plans include multiple PTVs and many individual salivary and swallowing OARs, testing RapidPlan performance in a relatively challenging, although common, clinical scenario. Pooled results showed that in general, if the majority of OARs in a patient were not flagged as outliers, RapidPlan provided comparable and often improved plan quality over the CPs. The superiority of certain RapidPlan plans over the CP could be due to the challenging nature of optimally and consistently performing interactive planning for plans which contain many OARs, within a limited number of iterations

The pooled data illustrates that models created using 30 plans created similar plans as a model based on 60 plans. However, these models could produce substantially different results in individual patients for specific OARs. This indicates that plans produced by the models are sensitive to the composition of the plan library and to the characteristics of the patient for which the knowledge-based plan is made. We also observed that more OAR outliers did not necessarily translate into a worse OAR dose.

RapidPlan results for specific OARs in EG1 showed appreciable variation between the models. For example, $Model_{30B}$ plans resulted on average in 3.2Gy / 2.8Gy higher D_{OC} than

$Model_{30A}$ / $Model_{60}$. The following factors may have contributed: (1) $Model_{30B}$ included the least oral cavities (24 versus 27 / 51 in $Model_{30A}$ / $Model_{60}$), (2) the oral cavity volumes in $Model_{30B}$ (mean 108cm^3, range 29-284cm^3) were on average larger than in EG1 (mean 87cm^3, range 35-242cm^3), (3) in the three patients that showed the largest increase in D_{OC} with $Model_{30B}$ compared to the CPs, one had a large PTV_B (266.5cm^3) and combined PTV volume (1044.1cm^3) compared to the model (averages of 150.4±102.1cm^3 and 594.7±222.4cm^3, respectively), The second had a small oral cavity (41.3cm^3, versus 108.2cm^3 in the model) and the third had both a small oral cavity (39.7cm^3) and a large PTV_B volume (666.7cm^3).

While it is expected that including more patients to the model libraries would lead to fewer outliers and more consistent plan quality, this requires further investigation. We observed for example that, although $Model_{60}$ resulted in fewer OAR outliers in EG1 and EG2, this did not translate into consistently lower OAR doses. For this study, patients were randomly selected for inclusion in the models, and no CPs were re-optimized on the basis of model statistics. The influence of the dosimetric outliers in $Model_{60}$ on resulting plan quality was evaluated by removing these outliers from the model, and replanning the first five patients of EG1. The effect was marginal, with HI_B / HI_E differing by less than 0.0±0.4% / 0.4±0.3%, on average, while D_{OC}, D_{sal} and D_{swal} differed by no more than 0.2±0.6%. Because RapidPlan results depend on the OAR-PTV geometry of the plans included in the models, consistent contouring is important in generating reliable model plan libraries. This might be helped by the use of contouring atlases[24]. Note that clinical inspection of plans created using RapidPlan remains necessary.

With respect to EG2, the RapidPlan models achieved substantial gains in sparing for most OARs over the CPs. These results indicate the gains in OAR sparing that inexperienced centers starting VMAT treatments might anticipate when using a RapidPlan model created by a more experienced center producing higher quality plans. They also show that a model comprising plans with many OARs is versatile and can be successfully applied to patients with fewer OARs.

Some of the improvements (e.g. in PTV dose homogeneity) can be attributed to the newer optimization algorithms available in Eclipse v13.5. This may be because the CPO functionality that uses the final, AAA calculated dose distribution as a starting point for a new optimization was not available in PRO v8.6-8.9 while it has been shown to substantially improve PTV dose homogeneity[19].

Results of several in-house developed knowledge-based planning solutions have been previously reported. Moore et al.[15] evaluated the correlation between the fraction of OARs

overlapping with the PTVs and the mean OAR doses and were able to identify patients in whom large gains in parotid gland sparing could be achieved. These gains were validated by re-planning the identified patients. Predicting achievable DVHs using a library of previously delivered plans has been investigated by various groups. Yuan et al.[16] used principal component analysis to identify significant patient anatomical factors contributing to OAR sparing in head and neck IMRT plans and applied this knowledge to realize gains in parotid gland sparing. Appenzoller et al.[14] identified differences between predicted and obtained DVHs to identify possible outliers with regards to OAR sparing. Replanning these outliers allowed refinement of the model.

In conclusion, RapidPlan knowledge-based treatment planning can deliver at least comparable results to CPs when the patient is similar to the bulk of patients used to populate the model. Caution should be used when applying RapidPlan models to patients whose geometry falls outside the range of the constituent plans in the model, in line with the advice contained in the instruction manual[22,23]. Determining whether an individual plan produced by RapidPlan is sub-optimal, defining when outlier detection should lead to plan rejection or manual planning and determining the optimal composition of model libraries, including the relationship between model composition and dosimetry of subsequent plans, requires further investigation.

References

1 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

2 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

3 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

4 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

5 T.S. Hong, D.L. Craft, F. Carlsson, and T.R. Bortfeld, "Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head.," Int. J. Radiat. Oncol. Biol. Phys. **72**(4), 1208–14 (2008).

6 D.L. Craft, T.S. Hong, H.A. Shih, and T.R. Bortfeld, "Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **82**(1), e83–90 (2012).

7 H. Chen, D.L. Craft, and D.P. Gierga, "Multicriteria optimization informed VMAT planning.," Med. Dosim. **39**(1), 64–73 (2014).

8 B. Wu, D. Pang, P. Simari, R. Taylor, G. Sanguineti, and T. McNutt, "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study.," Med. Phys. **40**(2), 021714 (2013).

9 Y. Yang, E.C. Ford, B. Wu, M. Pinkawa, B. van Triest, P. Campbell, D.Y. Song, and T.R. McNutt, "An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection.," Med. Phys. **40**(1), 011709 (2013).

10 B. Wu, T. McNutt, M. Zahurak, P. Simari, D. Pang, R. Taylor, and G. Sanguineti, "Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **84**(5), e647–53 (2012).

11 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

12 X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

13 V. Chanyavanich, S.K. Das, W.R. Lee, and J.Y. Lo, "Knowledge-based IMRT treatment planning for prostate cancer," Med. Phys. **38**(5), 2515 (2011).

14 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

15 K.L. Moore, R.S. Brame, D.A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **81**(2), 545–51 (2011).

16 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

17 M. Zarepisheh, T. Long, N. Li, Z. Tian, H.E. Romeijn, X. Jia, and S.B. Jiang, "A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning.," Med. Phys. **41**(6), 061711 (2014).

18 D. Good, J. Lo, W.R. Lee, Q.J. Wu, F.F. Yin, and S.K. Das, "A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers:

an example application to prostate cancer planning.," Int. J. Radiat. Oncol. Biol. Phys. **87**(1), 176–81 (2013).

19 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

20 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

21 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

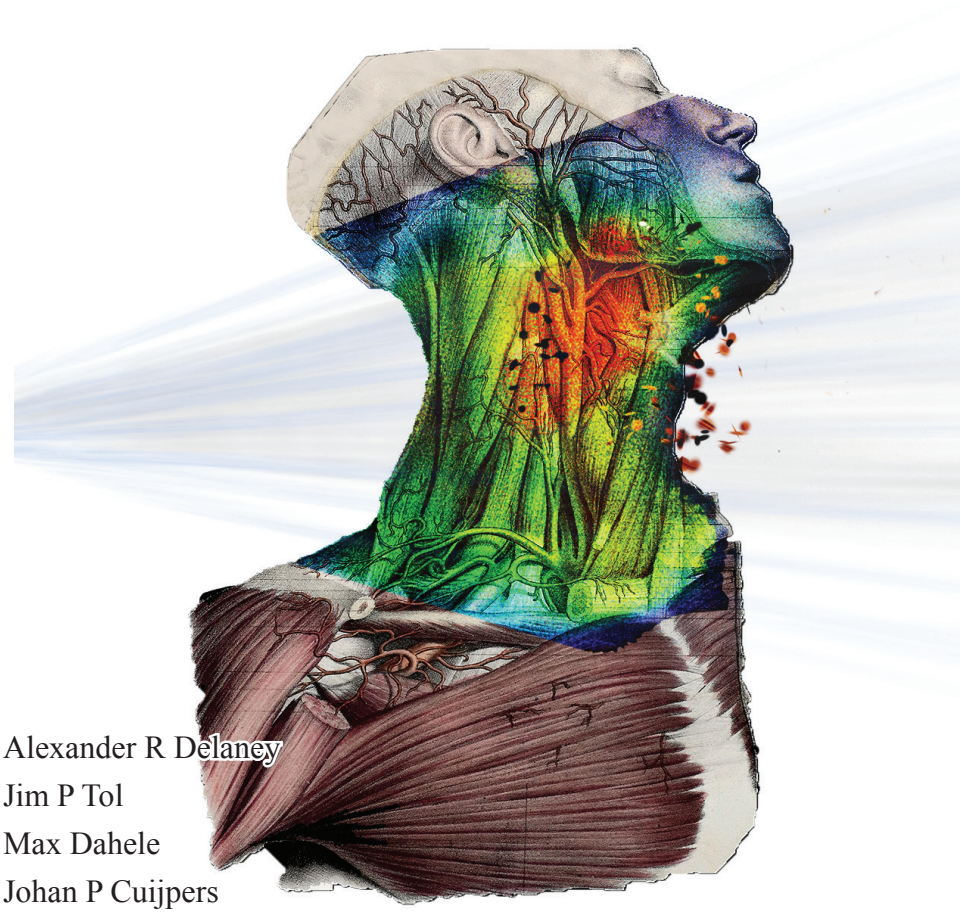
22 Varian Medical Systems, *Eclipse Photon and Electron Instructions for Use* (2014).

23 Varian Medical Systems, *Eclipse Photon and Electron Reference Guide* (2014).

24 M. Awan, J. Kalpathy-Cramer, G.B. Gunn, B.M. Beadle, A.S. Garden, J. Phan, E. Holliday, W.E. Jones, E. Maani, A. Patel, J. Choi, V. Clyburn, B. Tantiwongkosi, D.I. Rosenthal, and C.D. Fuller, "Prospective assessment of an atlas-based intervention combined with real-time software feedback in contouring lymph node levels and organs-at-risk in the head and neck: Quantitative assessment of conformance to expert delineation.," Pract. Radiat. Oncol. **3**(3), 186–93 (n.d.).

Chapter 8

Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution



Alexander R Delaney
Jim P Tol
Max Dahele
Johan P Cuijpers
Ben J Slotman
Wilko FAR Verbakel

International Journal of Radiation Oncology, Biology, Physics **94**(3), 469-477 (2016).

Abstract

Purpose

RapidPlan™, a commercial knowledge-based planning solution, uses a model library containing the geometry and associated dosimetry of existing plans. This model predicts achievable dosimetry for prospective patients which can be used to guide plan optimization. However, it is unknown how sub-optimal model plans (outliers) influence the predictions or resulting plans. We investigated the effect of firstly removing outliers from the model (cleaning) and subsequently adding deliberate dosimetric outliers.

Materials and Methods

Clinical plans from 70 head and neck cancer patients comprised the uncleaned $Model_{UC}$, from which outliers were cleaned to create $Model_{C}$. The last 5-40 patients of $Model_{C}$ were re-planned with no attempt to spare the salivary glands. These substantial dosimetric outliers were re-introduced to the model in increments of 5, creating $Model_{5\text{-}40}$. These models were used to create plans for a 10 patient evaluation group. Plans from $Model_{UC}$ and $Model_{C}$, and $Model_{C}$ and $Model_{5\text{-}40}$ were compared on boost / elective planning target volume homogeneity index (HI_{B} / HI_{E}) and mean dose to oral cavity, salivary glands ($comp_{sal}$) and swallowing structures ($comp_{swal}$).

Results

On average, outlier removal ($Model_{C}$ vs $Model_{UC}$) had minimal effects on HI_{B} / $HI_{E,}$ (0-0.4%), and OAR sparing, (mean dose difference to oral cavity / $comp_{sal}$ / $comp_{swal}$ ≤ 0.4Gy). $Model_{5/10}$ marginally improved $comp_{sal}$ sparing while adding a larger number of outliers ($Model_{20\text{-}40}$) led to deteriorations to $comp_{sal}$ mean dose of up to 3.9Gy, on average. These increases are modest compared to the 14.9Gy dose increases in the added outlier plans, due to the placement of optimization objectives below the inferior boundary of the DVH-prediction range.

Conclusions

Overall, dosimetric outlier removal from or adding 5-10 outliers to a 70-patient RapidPlan model had marginal effects on resulting plan quality. Although the addition of >20 outliers deteriorated plan quality, the effect was modest. In this study, RapidPlan demonstrated robustness for moderate proportions of salivary gland dosimetric outliers.

Introduction

Automated solutions to inverse radiotherapy treatment planning are being developed to minimize the variability in plan quality associated with manual planning and increase planning efficiency[1,2]. RapidPlan™ (Varian Medical Systems, Palo Alto, USA), a commercial knowledge-based planning solution[3–11] derived from earlier work[12,13], utilizes a model based on a library of previous plans. Regression analysis is used to discern correlations between the geometric and dosimetric features of the planning target volumes (PTVs) and organs-at-risk (OARs) of the library plans. The model can be used to predict a range of achievable OAR dose-volume histograms (DVHs) for new patients. RapidPlan (RP) subsequently guides the intensity modulated radiotherapy (IMRT) or RapidArc™ (Varian Medical Systems) volumetric modulated arc therapy (VMAT) optimization process in the Eclipse treatment planning system by placing a line of optimization objectives along the inferior boundary of the DVH-prediction range. The broadness of the prediction range, and hence the optimization objective placement, is influenced by the goodness of the fit between the geometric and dosimetric attributes of the model, along with the similarity of a prospective patient to the model population. Although knowledge-based planning using RP is in its infancy, promising results have been reported for different treatment sites[14–16].

RP performs statistical analysis on all modeled structures, thereby allowing the identification of OARs which deviate from the bulk of the model population. Because of concerns that such ´outliers´ reduce the goodness of fit between geometry and dosimetry[9], and could in turn negatively influence model performance, removal or replanning of outlier OARs is recommended by the vendor[17]. This however, is a time consuming and subjective process that must be repeated for each newly created or modified model. It is therefore useful to investigate the impact of this step on the created models. Although previous work illustrated the importance of having a broad range of OAR geometries in a model to accommodate prospective patients with varying geometries[16], the effect of dosimetric outliers in the model library on RP performance has not been systematically investigated.

The present study investigates the effect of dosimetric outliers on the resulting quality of generated RP plans for complex head and neck cancer (HNC) patients. Dosimetric and geometric outliers were removed from the model library and the effect on resulting plan quality of deliberately contaminating the model with plans in which the salivary glands were not actively spared was investigated.

Materials and Methods

All created models were populated with HNC RapidArc plans, planned with a simultaneous integrated boost technique using two full arcs and 6MV photons. Prescribed doses of 70Gy / 54.25Gy to the boost / elective PTV (PTV_B / PTV_E) were delivered in 35 fractions. The aim was to deliver 95% of the prescribed dose to 99% / 98% of PTV_B / PTV_E, while limiting the volume of each PTV receiving >107% of the prescribed dose. A 5mm transition region (PTV_T) was created between the PTVs to allow for a dose fall-off between them. The plans included sparing of the oral cavity (OC), salivary glands and swallowing muscles (delineated according to Christianen et al.[18]), although individual structures could be sacrificed by the treating clinician depending on the degree of overlap with the PTVs. Similar to our clinical optimization approach, the brainstem, spinal cord and their planning-at-risk volumes (3mm expansion) were assigned single maximum point dose objectives placed below their respective dose tolerance levels. Optimization was performed using the progressive resolution optimizer v10.0.28, followed by a dose calculation using the anisotropic analytical algorithm (AAA) v10.0.28. Subsequently, a continue previous optimization was performed for all plans to improve PTV dose homogeneity[19]. Previous work outlines the departmental approach to plan optimization in detail. In brief, parallel OARs were assigned 3 or 4 optimization objectives, positioned evenly along the DVH-line. During the interactive optimization process in Eclipse, the planner aimed to re-position these objectives to maintain approximately a fixed diagonal distance to the DVH[19].

Models and model cleaning

70 HNC patients were used to populate the initial model, $Model_{UC}$. Regression, residual and DVH-plots, combined with statistical metrics of the model data (outlined in previous work[14]) provided in the model configuration window, assist the user in outlier detection. Substantial dosimetric and / or geometric outliers were removed from $Model_{UC}$ in two iterations to create the 'cleaned' model, $Model_C$. Structures with DVHs positioned below the prediction range were considered positive dosimetric outliers, as the OAR was better spared than predicted by RP, and kept in the plan library. Structures with DVHs above the generated prediction range were deemed negative dosimetric outliers, meaning improved sparing could be achieved, and therefore removed from the library. Negative dosimetric / geometric outliers were furthermore removed in two iterations of cleaning by (1) identifying OARs which exceeded the threshold values of the statistical metrics; and (2) confirming these outliers through visual inspection of the regression, residual and geometric plots. Removal of geometric outliers may result in a lack of representation of

certain geometric features in the model library, which may in turn affect resulting DVH-predictions and dosimetry. Replanning rather than removing such outliers can retain the geometric features in the model. This however, is a time consuming process and since the bulk of $Model_C$ maintained considerable geometric variation, outlier replanning was not considered necessary (Figure 1).

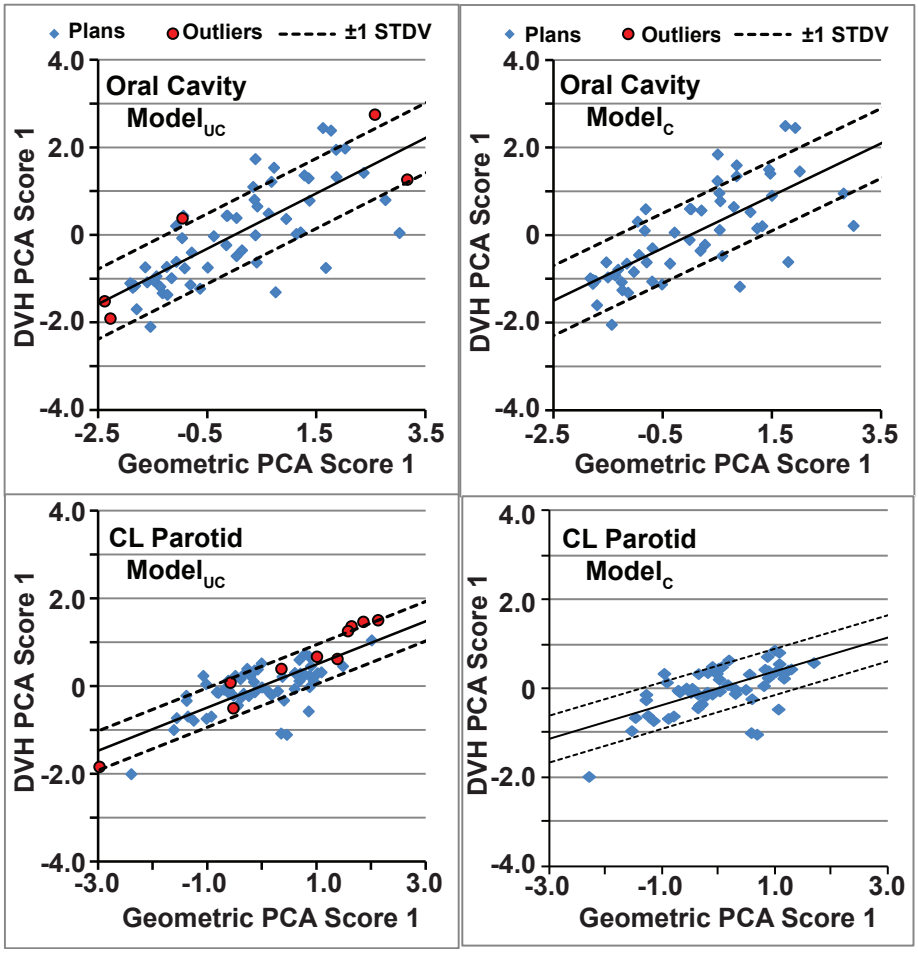


Figure 1. Regression plots (created using Model Analytics) for the oral cavity and contralateral parotid gland, using $Model_{UC}$ and $Model_C$. The red points indicate points which were deemed as outliers in model $Model_{UC}$ after analysis of the metrics in the model configuration window and visual interpretation of the geometric, regression and DVH plots. 10 / 5 outliers were removed from the contralateral parotid gland / oral cavity model structures in two iterations of cleaning, with $Model_C$ retaining 59 / 56 matched structures.

Contaminating models with dosimetric outliers

Generally, RP plan libraries will contain some dosimetric outliers where OARs are not optimally spared. To systematically investigate the effect of such outliers on RP performance, deliberate dosimetric outliers were introduced to $Model_C$. These outliers were made by replanning the last 5-40 patients (in increments of 5) of $Model_C$ without attempting to spare the salivary glands. This allowed examination into the number of outliers which would cause the model performance to be negatively influenced.

If few but substantial dosimetric outliers would only marginally influence model performance, it suggests that minor dosimetric deviations present in models containing clinical treatment plans would have a negligible impact on resulting plan quality. Salivary glands were used for this analysis because the contralateral / ipsilateral parotid and contralateral submandibular glands were well spared in most of the patients included in the model library. The majority of these deliberate outliers fell within the geometric range of the remaining model library. The plans containing the dosimetric outliers were added to $Model_C$ rather than $Model_{UC}$ to that ensure any subsequent effect on resulting plan quality would be a result of outlier contamination. This facilitated systematic investigation into the effect of varying numbers of extreme outliers and resulted in 8 new "outlier" models, $Model_5$-$Model_{40}$ (5 plan increments).

Evaluation group

The models were used to create RapidArc plans for a 10 patient evaluation group. RP plans were generated using the photon optimizer (PO) algorithm and AAA v13.5.33 using the same set of optimization objective weightings for all OARs. The generated plans were normalized to deliver the same PTV_B mean dose as their respective clinical plans. Composite salivary ($comp_{sal}$) and swallowing ($comp_{swal}$) structures, consisting of individually optimized OARs, were created to simplify OAR dose reporting. Table 1 shows the volumes of the PTVs, OC, $comp_{sal}$ and $comp_{swal}$ included in the models and evaluation group.

The RP plans for the evaluation group were compared on the basis of (i) PTV_B / PTV_E homogeneity index values (HI_B / HI_E), calculated using $HI = 100\% \times (D2\% - D98\%) / 50\%$, (ii) mean doses to individual and composite OARs, and (iii) differences between DVH-prediction ranges. Paired two-sided Student t-tests were performed to determine whether differences in obtained dosimetry were statistically significant ($p \leq 0.05$).

In summary, 10 models were used to generate RP plans for 10 evaluation patients. A comparison was made between plans resulting from (1) the uncleaned ($Model_{UC}$) and the

cleaned ($Model_C$) model, and (2) the outlier models ($Model_5$-$Model_{40}$) and the cleaned model ($Model_C$).

Table 1. Size of the planning target volumes (PTVs), oral cavities and composite salivary / swallowing structures ($comp_{sal}$ / $comp_{swal}$) for the 70 patients included in the RapidPlan model libraries and the 10 patient evaluation group.

	Models (n=70)		Evaluation Group (n=10)	
Volumes (cm^3)	**Mean ± StDev**	**Range**	**Mean ± StDev**	**Range**
PTV_B [a]	184.1 ± 113.2	34.1 to 607.0	194.3 ± 116.4	32.4 to 370.1
PTV_E [b]	368.2 ± 90.6	210.0 to 618.6	376.3 ± 78.1	283.9 to 537.3
PTV_T [c]	69.8 ± 42.7	10.1 to 258.4	55.8 ± 26.6	15.7 to 106.6
Oral Cavity	88.5 ± 57.6	14.7 to 283.5	90.9 ± 58.6	23.7 to 175.7
$Comp_{sal}$ [d]	63.0 ± 23.3	16.6 to 118.2	70.8 ± 12.8	39.7 to 83.7
$Comp_{swal}$ [e]	22.7 ± 13.5	1.3 to 62.5	25.9 ± 17.1	7.0 to 58.2

[a] Boost planning target volume
[b] Elective planning target volume
[c] Transition planning target volume
[d] Composite salivary structures
[e] Composite swallowing structures

Results

Outlier removal

In the first iteration of the model cleaning process, 76 OARs were identified with at least one of the statistical metrics exceeding the threshold values. Of these, 28 were visually confirmed to be outliers including 22 and 6 dosimetric and geometric outliers, respectively. In the second iteration, 62 OARs exceeded the thresholds and 19 were visually confirmed to be outliers, containing 7 and 12 dosimetric and geometric outliers, respectively. The majority of dosimetric outliers were identified in the first iteration, with mean dose differences between the achieved and predicted OAR DVHs of 6.3Gy, on average. Geometric outliers were mostly identified in the second iteration, with average mean dose differences of 3.1Gy. Table 2 shows examples of the model populations.

Outlier removal marginally affected resulting dosimetry with changes of <0.4Gy, on average, to the OC, $comp_{sal}$ and $comp_{swal}$ mean doses (Table 3). In one patient however, the cleaning process resulted in substantial dosimetric differences. The mean dose to the upper larynx of patient 8 increased from 35.1Gy to 54.1Gy while mean dose to the medial pharyngeal constrictor muscle (PCM) increased from 38.8Gy to 50.9Gy. These increases were caused by the removal of structures with similar target and OAR overlap volumes from $Model_C$ in

the cleaning process, resulting in poor modeling of the patient's upper larynx and PCM. Compared to $Model_{UC}$, 5 individual improvements and degradations greater than 2Gy were obtained using $Model_{C}$.

Table 2. The number of matched structures to each organ-at-risk (OAR) model structure in $Model_{UC}$ and $Model_{C}$. The geometry is well represented after the outlier removal process (Figure 1) and this can be attributed to the small number of structures removed during the outlier removal process to create $Model_{C}$.

OAR model structure	$Model_{UC}$	$Model_{C}$
CL Parotid[a]	69	59
CL SMG[b]	51	48
IL Parotid[a]	67	67
Oral Cavity	61	56
Cricoph[c]	54	50
Lower Larynx	45	42
Upper Larynx	32	24
Inferior PCM[d]	42	38
Medial PCM[d]	30	25
Superior PCM[d]	37	36
UES[e]	69	65

[a] Contralateral and ipsilateral parotid glands

[b] Contralateral submandibular gland

[c] Cricopharyngeal muscle

[d] Pharyngeal constrictor muscle

[e] Upper esophageal sphincter

Contaminating models with dosimetric outliers

Mean $comp_{sal}$ dose in the 40 outlier plans was on average 14.9Gy higher than their respective original plans. The contralateral parotid (CLP) regression plots (Figure 2) clearly show the deviation of these outliers from the bulk of original plans, with a progressive widening in the standard deviation (SD) of the fit as more outliers were added.

The decreased correlation between geometry and associated dosimetry in the outlier models resulted in a broadening of the DVH-prediction range for the salivary glands, and an upwards shift of the mean DVH-prediction (Figure 3). In models with 5-10 outliers ($Model_{5}$-$Model_{10}$), the inferior boundary of the DVH-prediction moved to lower doses. Since RP places optimization objectives along this boundary, this resulted in marginal improvements in OAR sparing using these outlier models, on average. For 8 out of 10

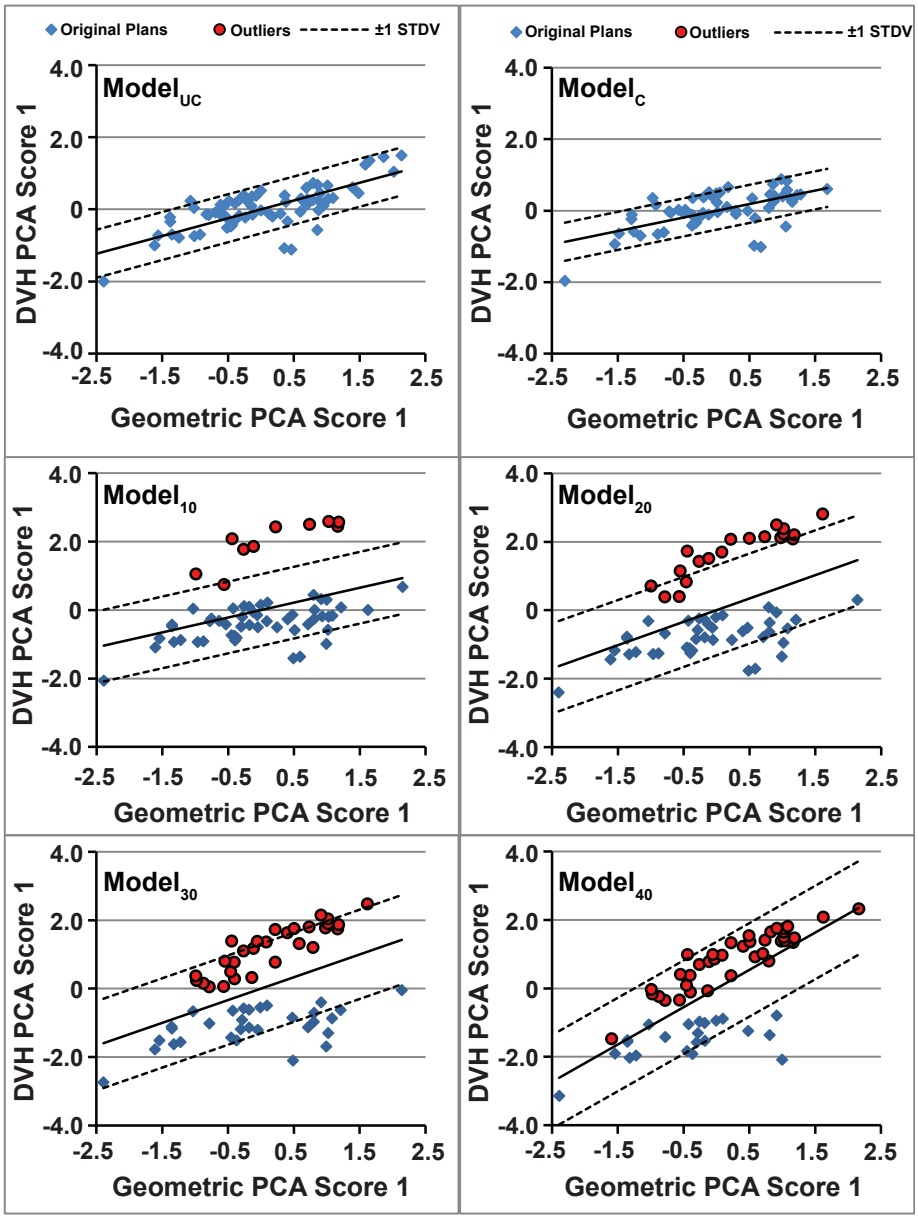


Figure 2. Regression plots (created using Model Analytics) for the contralateral parotid gland using Model$_{UC/C/10/20/30/40}$. These plots show the correlation between the most prominent principal components of the geometric (x-axis) and dosimetric (y-axis) features. The first principal component can correspond to an array of features, depending on the contribution of certain features to the explanation of the variance. While removing outliers from Model$_{UC}$ to create Model$_{C}$ does not have a substantial impact, the addition of outliers in Model$_{10/20/30/40}$ (indicated in red) resulted in a progressive increase in the standard deviation with a concurrent change in the slope.

CLPs and 7 out of 9 ipsilateral parotids, both $Model_5$ and $Model_{10}$ reduced mean OAR doses (>2Gy in 3 cases). As the number of outliers increased, however, the lower boundary typically shifted upwards, with evident increases in salivary doses for models with ≥20 outliers (Figure 4, Table 3). In patients 1 and 3 (Figure 5), for example, $Model_{40}$ shifted the lower prediction boundary for the CLP upwards considerably, resulting in an 8.0Gy and 10.0Gy increase over $Model_C$, respectively. Similar results were seen for the contralateral submandibular gland. In a few patients, however, $Model_{40}$ still performed well. In patient 2, for example, the upwards shift of the lower boundary was minimal, with a mean dose increase to the CLP of only 1.2Gy using $Model_{40}$ over $Model_C$.

Figure 4 shows that patients with such small mean dose increases typically clustered at lower geometric principal component values in the regression plots of both $Model_C$ and $Model_{40}$. In patient 2, for example, the upwards shift of the lower boundary was minimal, which is reflected in the mean dose increase to the CLP of only 1.2Gy using $Model_{40}$ over $Model_C$. Figure 6 shows that patients with such small mean dose increases using $Model_{40}$ were typically found to be clustered at lower geometric principal component values in the regression plots (closed circles). The bottom graph furthermore shows that, dosimetrically, all evaluation patients are located amongst the lower portion of the model population (i.e. below the solid line), consistent with the fact that the mean CLP doses of the evaluation patients using $Model_{40}$ increased by only 4.4Gy, on average, over $Model_C$ plans. Meanwhile, mean doses of the 40 CLP outliers were 14.5Gy higher, on average, compared to their respective clinical plans and resultantly lie above the solid line. The relatively small upward shift of the inferior DVH-prediction range using $Model_{40}$ (Figure 3) limits the influence of these outliers on the resulting plans, leading to modest deteriorations overall.

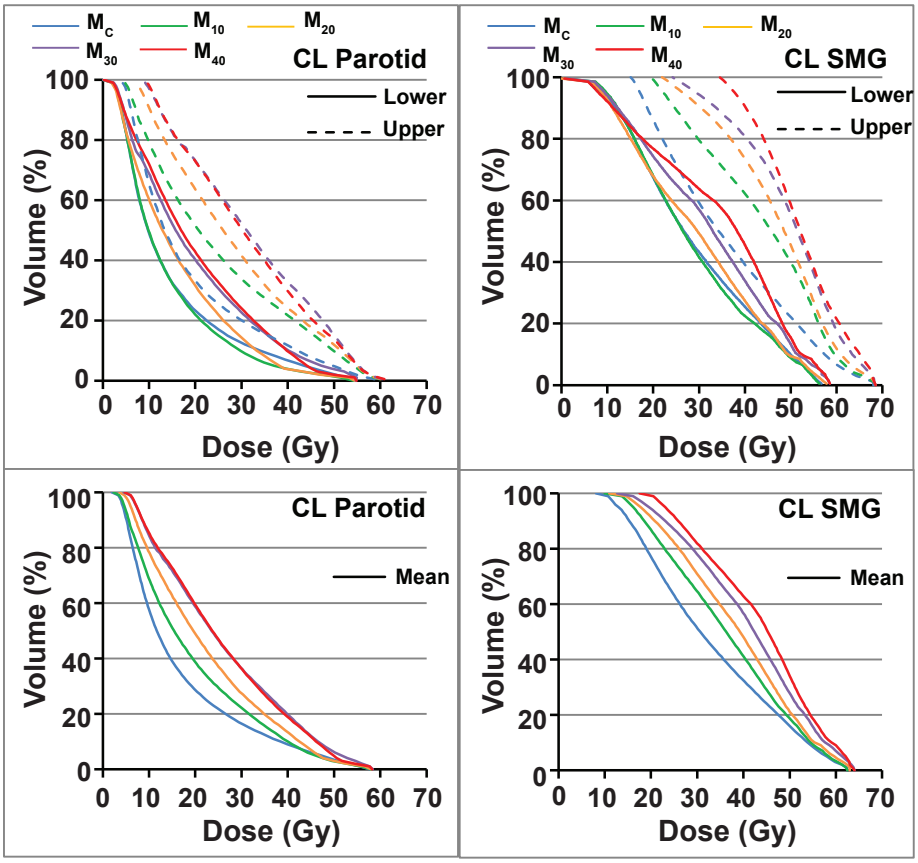


Figure 3. Averaged lower & upper (upper graphs) and mean (lower graphs) lines of the dose-volume histogram (DVH)-prediction range for the contralateral parotid (CL Parotid) and submandibular (CL SMG) glands using $Model_C$ (blue), $Model_{10}$ (green), $Model_{20}$ (orange), $Model_{30}$ (purple), and $Model_{40}$ (red).

Table 3. RapidPlan results for the 10 patient evaluation group using the initial model ($Model_{UC}$), cleaned model ($Model_{C}$) and outlier models ($Model_{5\text{-}40}$). Results are averaged over all 10 patients. Mean OAR doses are reported.

Model	$Model_{UC}$	$Model_{C}$	$Model_{5}$	$Model_{10}$	$Model_{15}$	$Model_{20}$	$Model_{25}$	$Model_{30}$	$Model_{35}$	$Model_{40}$
Boost planning target volume (%)										
PTV_{B} V95	99.3 ± 0.6	99.3 ± 0.5	99.2 ± 0.5	99.2 ± 0.6	99.1 ± 0.6	99.3 ± 0.5	99.3 ± 0.6	99.4 ± 0.5	99.4 ± 0.5	99.4 ± 0.5
PTV_{B} V107	2.2 ± 2.6	2.2 ± 2.9	2.2 ± 3.0	2.6 ± 3.3	2.8 ± 3.7	2.3 ± 2.7	1.8 ± 2.7	1.7 ± 2.7	2.0 ± 3.1	1.9 ± 2.5
HI[a]	9.4 ± 1.5	9.4 ± 1.3	9.4 ± 1.4	9.6 ± 1.6	9.8 ± 1.8	9.4 ± 1.3	9.1 ± 1.5	8.9 ± 1.3	9.1 ± 1.5	9.1 ± 1.2
Elective planning target volume (%)										
PTV_{E} V95	98.3 ± 0.8	98.3 ± 0.7	98.2 ± 0.9	98.2 ± 0.8	98.1 ± 0.9	98.2 ± 0.9	98.4 ± 0.8	98.4 ± 0.8	98.6 ± 0.5*	98.6 ± 0.7*
PTV_{E} V107	12.2 ± 5.8	11.3 ± 6.0	12.0 ± 6.6	11.3 ± 5.3	12.2 ± 6.3	11.4 ± 5.5	11.1 ± 5.5	10.5 ± 5.6*	10.9 ± 5.3	10.1 ± 5.6
HI[a]	14.7 ± 1.5	14.3 ± 1.1	14.5 ± 1.2	14.6 ± 1.6	14.8 ± 1.9	14.6 ± 1.4	14.3 ± 1.1	14.1 ± 1.3	13.9 ± 1.1*	13.5 ± 1.4*
Mean dose (Gy)										
Oral Cavity	20.5 ± 8.4	20.6 ± 9.0	20.3 ± 8.5	21.1 ± 9.7	20.1 ± 8.5	20.0 ± 8.5	20.0 ± 8.5	19.6 ± 8.4	19.6 ± 8.2	19.6 ± 8.2
CL Parotid[b]	16.8 ± 5.2	16.7 ± 5.2	16.3 ± 5.6	16.2 ± 5.5	18.0 ± 7.2	18.1 ± 7.3	18.6 ± 7.6	20.2 ± 7.7*	18.4 ± 6.5	21.1 ± 8.9*
IL Parotid[b]	19.9 ± 6.5	19.9 ± 6.5	19.2 ± 6.3*	19.0 ± 6.4	20.2 ± 7.4	20.4 ± 7.9	21.0 ± 7.8	21.1 ± 7.6	23.0 ± 7.7*	23.8 ± 8.8*
CL SMG[c]	35.0 ± 7.6	34.8 ± 8.2	34.6 ± 8.6	34.7 ± 8.9	34.7 ± 8.0	34.8 ± 8.4	34.6 ± 8.6	36.5 ± 9.2*	37.3 ± 8.5	38.0 ± 7.8*
$Comp_{sal}$[d]	21.4 ± 4.7	21.3 ± 4.6	20.8 ± 4.8*	20.7 ± 5.0*	22.0 ± 5.5	22.2 ± 5.8	22.7 ± 5.8	23.8 ± 5.8*	23.8 ± 5.1*	25.2 ± 6.9*
$Comp_{swal}$[d]	27.9 ± 6.0	28.3 ± 6.0	27.9 ± 6.2	27.9 ± 6.0	27.7 ± 6.1	27.6 ± 6.2	27.8 ± 6.0	27.5 ± 6.0*	27.4 ± 6.3*	27.4 ± 6.3

* Statistically significant difference with respect to the $Model_{C}$ plan

[a] Homogeneity index for the boost and elective planning target volumes

[b] Contralateral and ipsilateral parotid glands

[c] Contralateral submandibular gland

[d] Composite salivary and swallowing structures

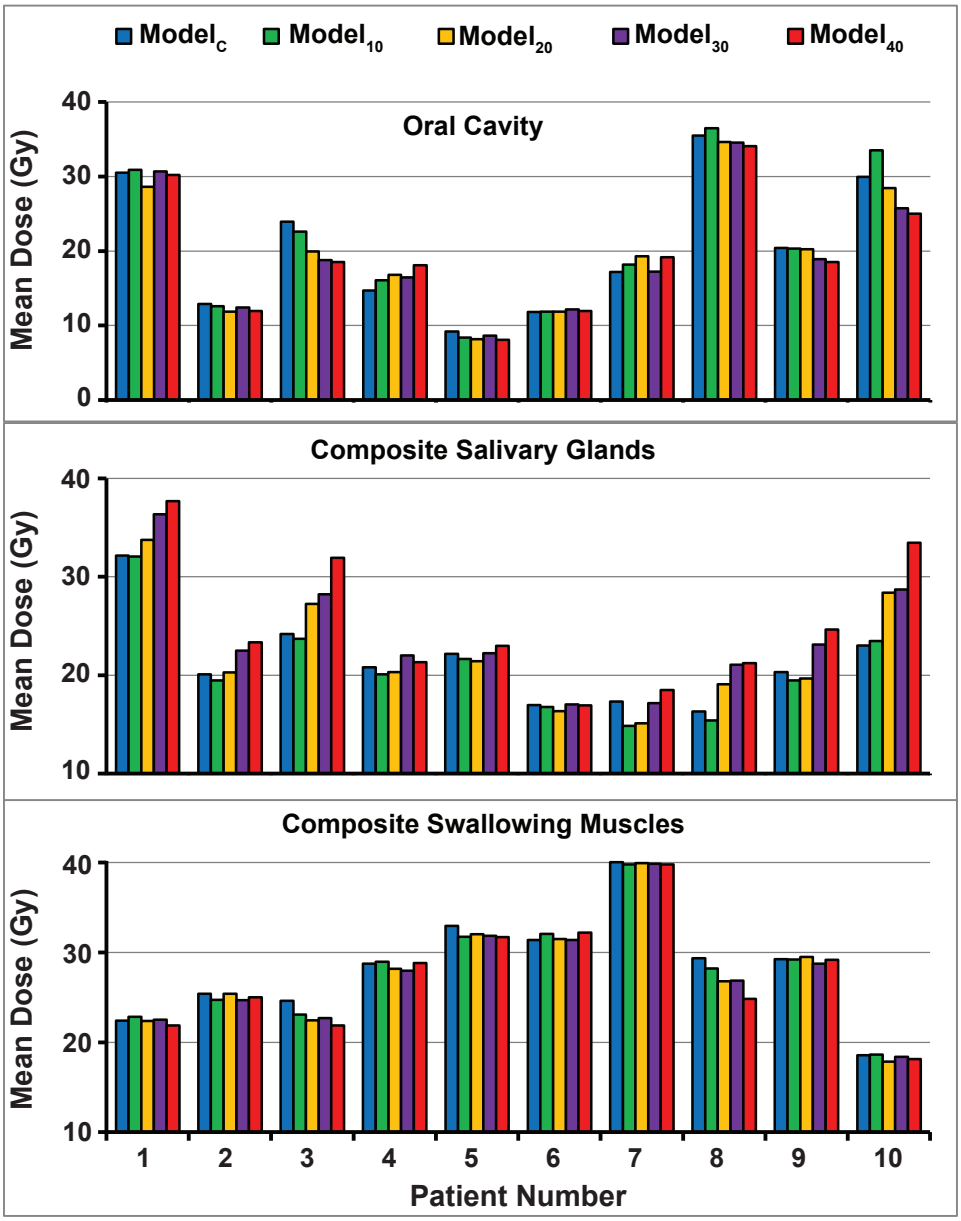


Figure 4. Mean dosimetric results for the 10 patient evaluation group using Model$_C$ (blue), Model$_{10}$ (green), Model$_{20}$ (orange), Model$_{30}$ (purple), and Model$_{40}$ (red).

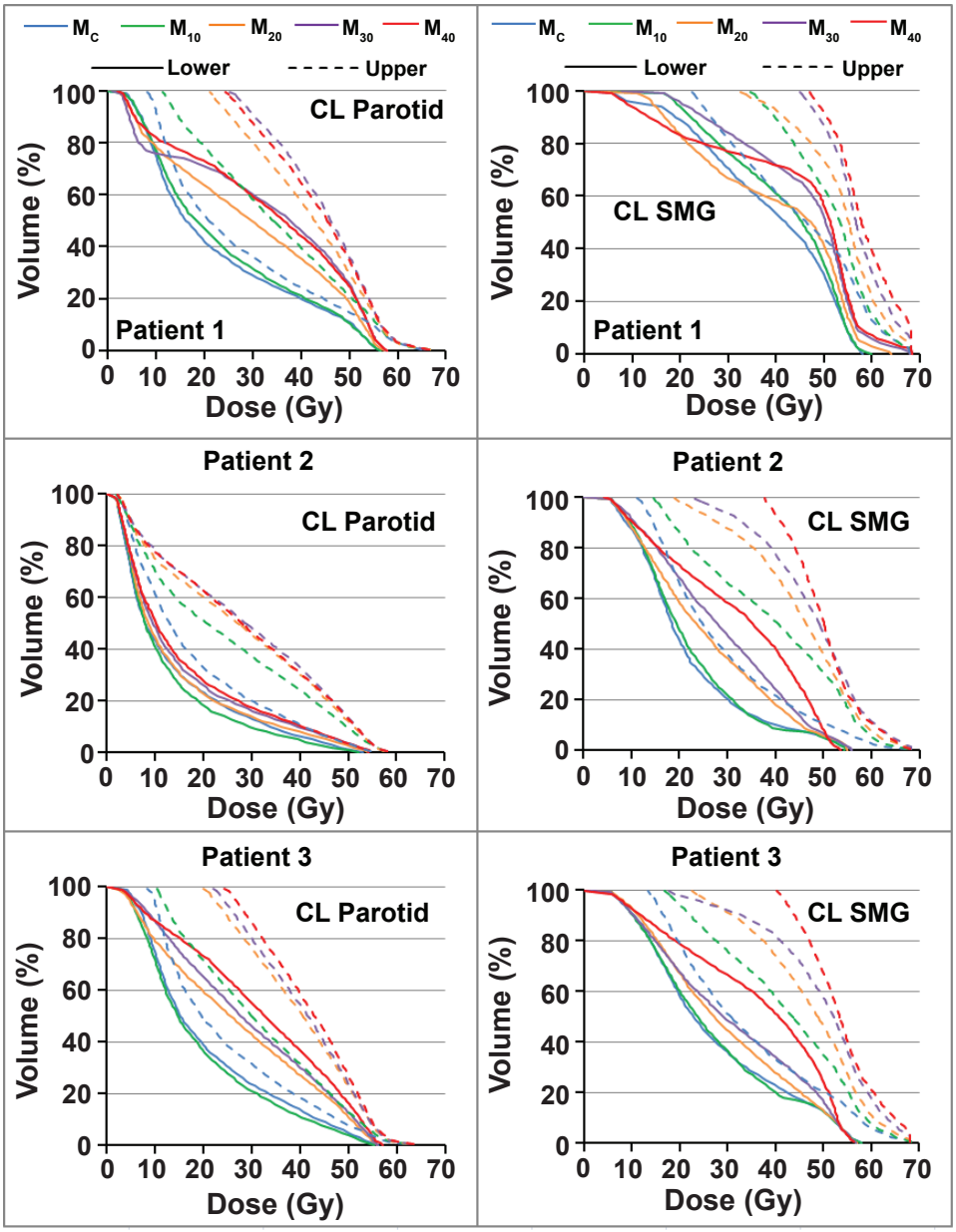


Figure 5. DVH-prediction ranges for 3 patients generated for the contralateral parotid and submandibular gland using $Model_C$ (blue), $Model_{10}$ (green), $Model_{20}$ (orange), $Model_{30}$ (purple) and $Model_{40}$ (red). Solid lines indicate the lower boundary while dashed lines indicate the upper boundary.

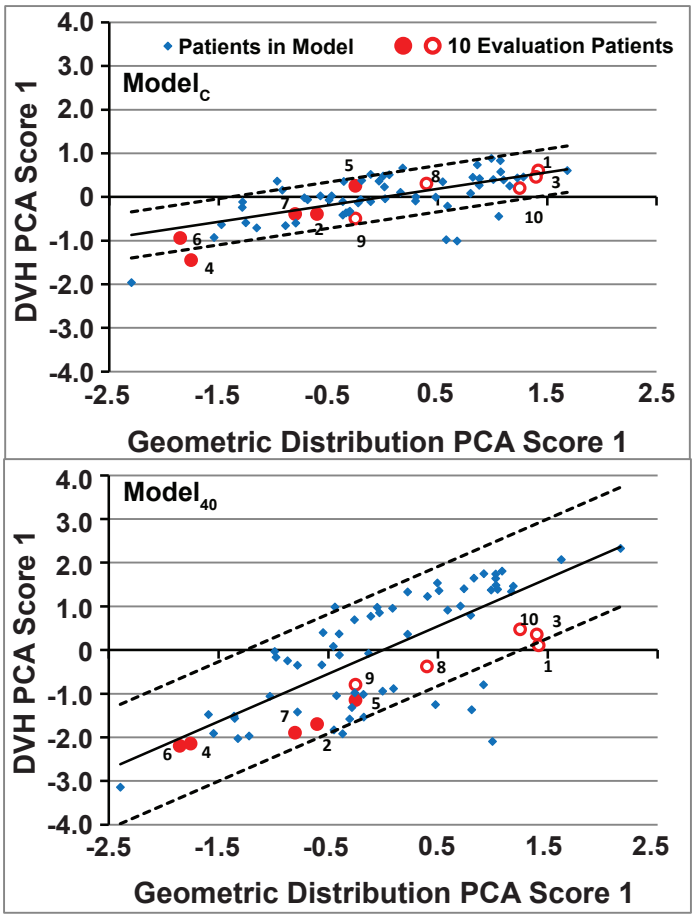


Figure 6. Regression plots for the contralateral parotid gland using Model$_C$ (upper graph) and Model$_{40}$ (lower graph), generated using Model Analytics (accessible to all users). Red filled circles indicate the evaluation patients which displayed ≤1.6Gy increase to the contralateral parotid gland mean dose, while open circles indicate patients in which greater deteriorations were noted, when using Model$_{40}$ over Model$_C$.

Discussion

This study revealed that removing geometric and dosimetric outliers from a RP model library did not generally improve resulting plan quality, questioning the necessity for such an extensive process. Deliberately adding dosimetric outliers to a 70 patient model library resulted in a broadening of the generated DVH-prediction range. For 5-15 outliers, averaged mean salivary gland doses were marginally affected. The quality of plans produced by the model was seen to deteriorate when more than 20 out of 70 patients were dosimetric outliers. $Model_{40}$ was noted to increase salivary gland mean doses by 3.9Gy, on average, even though the outlier plans added to the model contained salivary glands that were 14.9Gy higher, on average. Although obtained dosimetric values are dependent on the relative geometry of the dosimetric outliers with respect to the model library, the majority of a model should consist of good quality plans since larger numbers of dosimetric outliers deteriorates resulting plan quality. Progressive widening of the DVH-prediction ranges also prevents using the model for quality assure (QA) of the achieved DVHs of other plans[20].

It should be noted that the plans contained in the model library were all created in a relatively consistent manner, with standardized optimization objective weightings, planning aims and technique, independent of the treating radiation oncologist. This approach to planning may have contributed to the fact that only minor improvements in plan quality were found after cleaning the models. Other institutes may differ in their planning approach, with, for example, distinct planning aims varying between individual clinicians. This could introduce extensive variation into a model library, and as a result, the cleaning of outliers may result in greater improvements. We noted a number of cases with higher individual OAR doses using the cleaned model, which could be caused by the removal of OAR geometries in the cleaning process, which influenced the resulting predictions for similar OARs in the evaluation group[16]. This could be prevented by re-planning the patient rather than removing outlier OARs from the model.

Previous work illustrated promising results using knowledge-based planning (including RP) for a range of disease sites and with varying levels of patient complexity[3,10,11,14–16]. Fogliata et al.[14] retained potential dosimetric and geometric outliers identified by RP in the model. Visual inspection of these outliers indicated that there was no evidence of remarkably different cases and subsequently obtained clinically acceptable results using the model. The authors highlighted the need for further understanding the effect of outliers on resulting plans. Tol et al.[16] suggested the importance of including varying geometric features in the model but did not remove outliers from the model libraries. To the best of our knowledge,

the effect of outlier removal and deliberate inclusion of plans with inferior OAR sparing (dosimetric outliers) in a RP model on resulting plan quality has not been systematically investigated previously.

Visual analysis of the regression, residual and geometric plots of the model configuration window provided the most beneficial means of analyzing a model. Regression plots of $Model_{UC}$ and $Model_{C}$ (Figure 1) remained largely similar despite the outlier removal process, representative of the negligible differences in dosimetry (Table 3). Furthermore, adding dosimetric outliers to a model did not always violate the threshold values of the statistical metrics but visual analysis of the aforementioned plots allowed the user to identify which points were outliers (Figure 2). Model Analytics, an on-line tool available to RP users, can be used to further examine created models. By providing an overview of the mean doses and volumes of individual OARs, Model Analytics helps to investigate why certain points deviate from the dosimetric or geometric features of the model library.

A possible limitation of the present work is the standard set of optimization objective weightings (priorities) that was used for the PTVs and OARs. RP can calculate optimal objective priorities, aiming to take into account the accuracy of the DVH-prediction. In effect, a line objective for an OAR with a wider DVH-prediction range (thus larger SD in the regression) would be assigned a lower optimization priority to compensate for the uncertain prediction. This function could therefore, in theory, prevent lower mean dose values from being obtained using the models containing dosimetric outliers. However, initial evaluation of this function by our group did not result in better plans than using our own standard priorities and at the present time we have therefore not evaluated this further. Potential improvements to the priority generator requires future work to assess whether it can lead to plans with improved OAR sparing, and how it copes with outliers in the model. Furthermore, we only evaluated the performance of a HNC model that contained a large number of individual OARs, including many small swallowing muscles. These results may not be characteristic for other tumor sites. Since only salivary glands were used to create the outlier models, the DVH-prediction ranges for the individual salivary glands widened, while those of the OC and individual swallowing muscles hardly changed. This improved swallowing muscle and OC sparing using $Model_{20}$-$Model_{40}$ because less of an attempt was made to reduce salivary gland doses in these models. This trade-off would not be expected to occur if plans which poorly spared all OARs were added to the models. Finally, the majority of the 40 outliers added to the models were within the geometrical bounds of the remaining model population. If both geometric and dosimetric outliers would have been added this could have had a greater effect on the resulting dosimetry.

In conclusion, the present results show that removal of negative dosimetric outliers from a consistent HNC RP model did not improve the quality of plans made by the model. In addition, small gains in OAR sparing were found when 5 or 10 dosimetric outliers were included in a 70-plan model due to placement of the line objective along the inferior boundary of the DVH-prediction range by RP. For the same reason, the increase in salivary gland dose of the plans created by the model with over 20 outliers was modest compared to that of the added outlier plans. RP did not identify the majority of added outliers, highlighting the need for visual analysis of the model population.

References

1 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

2 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

3 B. Wu, D. Pang, P. Simari, R. Taylor, G. Sanguineti, and T. McNutt, "Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: a head-and-neck case study.," Med. Phys. **40**(2), 021714 (2013).

4 Y. Yang, E.C. Ford, B. Wu, M. Pinkawa, B. van Triest, P. Campbell, D.Y. Song, and T.R. McNutt, "An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection.," Med. Phys. **40**(1), 011709 (2013).

5 B. Wu, T. McNutt, M. Zahurak, P. Simari, D. Pang, R. Taylor, and G. Sanguineti, "Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **84**(5), e647–53 (2012).

6 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

7 X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

8 V. Chanyavanich, S.K. Das, W.R. Lee, and J.Y. Lo, "Knowledge-based IMRT treatment planning for prostate cancer," Med. Phys. **38**(5), 2515 (2011).

9 K.L. Moore, R.S. Brame, D.A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **81**(2), 545–51 (2011).

10 M. Zarepisheh, T. Long, N. Li, Z. Tian, H.E. Romeijn, X. Jia, and S.B. Jiang, "A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning.," Med. Phys. **41**(6), 061711 (2014).

11 D. Good, J. Lo, W.R. Lee, Q.J. Wu, F.F. Yin, and S.K. Das, "A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning.," Int. J. Radiat. Oncol. Biol. Phys. **87**(1), 176–81 (2013).

12 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

13 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

14 A. Fogliata, P.M. Wang, F. Belosi, A. Clivio, G. Nicolini, E. Vanetti, and L. Cozzi, "Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer.," Radiat. Oncol. **9**(1), 236 (2014).

15 A. Fogliata, F. Belosi, A. Clivio, P. Navarria, G. Nicolini, M. Scorsetti, E. Vanetti, and L. Cozzi, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer.," Radiother. Oncol. **113**(3), 385–91 (2014).

16 J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Evaluation of a knowledge-based planning solution for head and neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **91**(3), 612–20 (2015).

17 Varian Medical Systems, *Eclipse Photon and Electron Reference Guide* (2014).

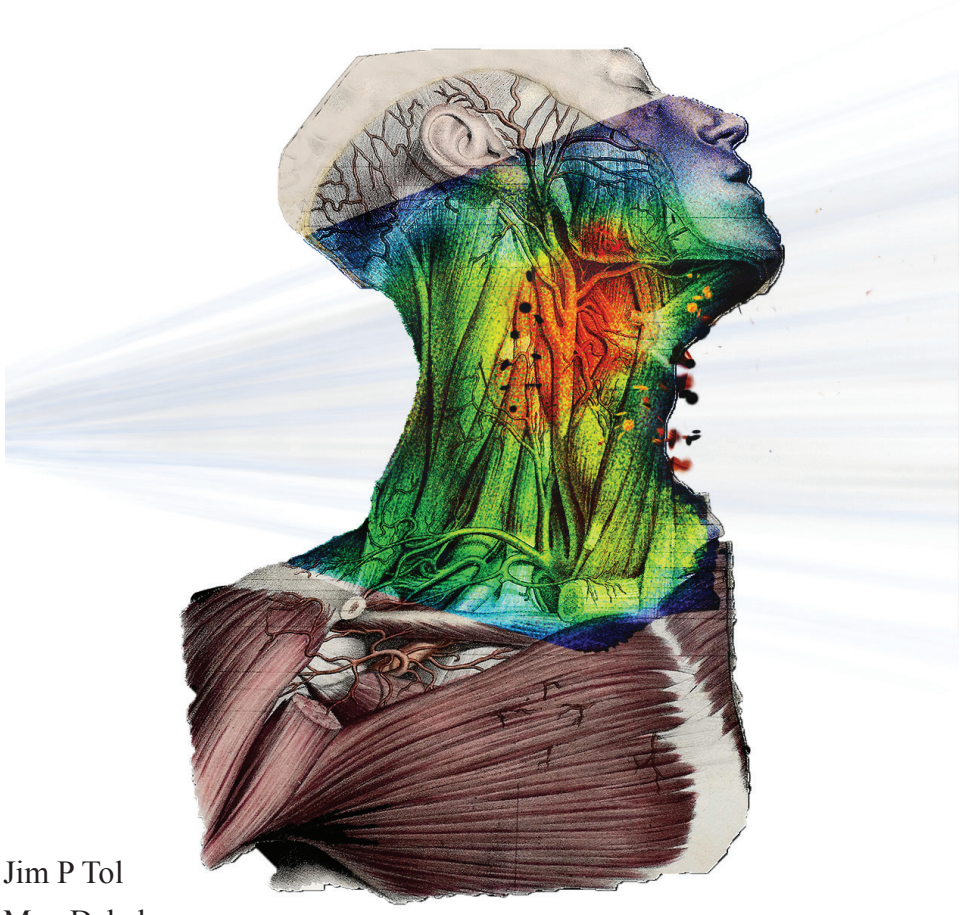
18 M.E.M.C. Christianen, J.A. Langendijk, H.E. Westerlaan, T.A. van de Water, and H.P. Bijl, "Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.," Radiother. Oncol. **101**(3), 394–402 (2011).

19 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

20 K.L. Moore, R. Schmidt, V. Moiseenko, L.A. Olsen, J. Tan, Y. Xiao, J. Galvin, S. Pugh, M.J. Seider, A.P. Dicker, W. Bosch, J. Michalski, and S. Mutic, "Quantifying Unnecessary Normal Tissue Complication Risks due to Suboptimal Planning: A Secondary Study of RTOG 0126.," Int. J. Radiat. Oncol. Biol. Phys. **92**(2), 228–35 (2015).

Chapter 9

Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans?



Jim P Tol
Max Dahele
Alexander R Delaney
Ben J Slotman
Wilko FAR Verbakel

Radiation Oncology **10**(1), 234 (2015).

Abstract

Purpose

Treatment plan quality assurance (QA) is important for clinical studies and for institutions aiming to generate near-optimal individualized treatment plans. However, determining how good a given plan is for that particular patient (individualized patient or plan QA, in contrast to running through a checklist of generic QA parameters applied to all patients) is difficult, time consuming and operator-dependent. We therefore evaluated the potential of RapidPlan, a commercial knowledge-based planning solution, to automate this process, by predicting achievable OAR doses for individual patients based on a model library consisting of historical plans with a range of organ-at-risk (OAR) to planning target volume (PTV) geometries and dosimetries.

Materials and Methods

A 90-plan RapidPlan model, generated using previously created automatic interactively optimized (AIO) plans, was used to predict achievable OAR dose-volume histograms (DVHs) for the parotid glands, submandibular glands, individual swallowing muscles and oral cavities of 20 head and neck cancer (HNC) patients using a volumetric modulated (RapidArc) simultaneous integrated boost technique. Predicted mean OAR doses were compared with mean doses achieved when RapidPlan was used to make a new plan. Differences between the achieved and predicted DVH-lines were analyzed. Finally, RapidPlan predictions were used to evaluate achieved OAR sparing of AIO and manual interactively optimized plans.

Results

For all OARs, strong linear correlations (R^2 = 0.94-0.99) were found between predicted and achieved mean doses. RapidPlan generally overestimated the amount of achievable sparing for OARs with a large degree of OAR-PTV overlap. RapidPlan QA using predicted doses alone identified that for 50% (10 / 20) of the manually optimized plans, sparing of the composite salivary glands, oral cavity or composite swallowing muscles could be improved by at least 3Gy, 5Gy or 7Gy, respectively, while this was the case for 20% (4 / 20) AIO plans. These predicted gains were validated by replanning the identified patients using RapidPlan.

Conclusions

Strong correlations between predicted and achieved mean doses indicate that RapidPlan could accurately predict achievable mean doses. This shows the feasibility of using RapidPlan DVH predictions alone for automated individualized head and neck plan QA. This has applications in individual centers and clinical trials.

Introduction

The increasing complexity of radiotherapy treatment planning, particularly due to the attempt to spare more individual organs-at-risk (OARs), has made it challenging to efficiently produce consistent, high quality radiotherapy treatment plans[1,2]. This has led to considerable interest in (semi)-automated planning strategies[3–9]. Plan quality assurance (QA) is another step in the treatment preparation workflow that might benefit from increased automation, since often, insufficient attention is given to evaluating whether a given plan can be improved. The time consuming[10], difficult and subjective nature of this process QA makes it hard to be confident that good plan quality is being obtained for individual patients. Furthermore, sub-optimal plans submitted to clinical trials have been correlated with worse clinical outcomes[11], suggesting that high quality plans are important for maximizing treatment outcomes. However, the fact that sub-optimal plans were accepted to the trial identifies a clear need for a robust and efficient plan QA tool to determine in near real-time whether plans meet an acceptable quality standard for individual patients[12]. Although Moore et al.[13] have recently published an analysis in which the number of patients at unnecessary risk for normal tissue complication probability was determined by comparing the predicted and achieved dosimetry, they used an in-house developed knowledge-based planning solution[14].

In contrast we have investigated whether OAR dose-volume histograms (DVHs) predicted by RapidPlan™, a commercial knowledge-based planning solution by Varian Medical Systems (Palo Alto, USA), are accurate enough to serve as a plan QA tool that could be used in clinical departments or for clinical trials. RapidPlan utilizes a library of plans to construct a model[14–17]. This model uses the geometrical features and associated dosimetry of the plans included in the library to predict a range of achievable DVH-lines for OARs of new patients. This range consists of the mean estimated DVH-line ± one standard deviation. The accuracy of the DVH prediction is therefore influenced by the consistency of the plans and the range of patient geometries in the library. The DVH-lines obtained during treatment planning for a patient can be compared with the DVHs predicted by RapidPlan. If the RapidPlan predictions are accurate, this comparison can be used to determine whether the evaluated plan has achieved adequate OAR sparing, as judged against the model library. This approach does not require the creation of additional plans and it would therefore present a fast and straightforward solution for plan QA. Although there have been reports of improvements in OAR sparing using RapidPlan[5–7], these do not necessarily imply that there was a close correspondence between the achieved and predicted DVH-lines, nor has this relationship been evaluated.

Our work differs from previous publications[14,18] in several ways, including (i) using a commercial solution to predict achievable DVHs, (ii) the use of complex head and neck cancer (HNC) treatment plans, involving sparing of the parotid glands, submandibular glands, oral cavity and many individual swallowing muscles, along with two PTVs and a simultaneous integrated boost, (iii) a detailed evaluation of the relationship between predicted and achieved dosimetry along the entire DVH-curve, and (iv) demonstration of QA application by using DVH predictions generated by RapidPlan to benchmark previously created plans.

Materials and Methods

General description of treatment planning

All HNC plans were created with a simultaneous integrated boost (SIB) technique using 6MV photons and 2 full RapidArc™ (Varian Medical Systems, Palo Alto, USA) arcs. In 35 fractions, plans aimed to deliver 95% of the prescribed dose of 54.25Gy / 70.0Gy to 98% / 99% of the elective / boost PTV (PTV_E / PTV_B), while limiting the volume of each PTV receiving >107% of the prescribed dose. A 5mm transition zone (PTV_T) was created between PTV_E and PTV_B to facilitate dose fall-off between them. The optimization goals and included OARs have been outlined in detail previously[7,19–21].

Model Library

Ninety HNC patients treated between 2012 and 2014 were arbitrarily selected. Primary tumor locations included the oropharynx (n=47), (supra-)glottic larynx (n=25), hypopharynx (n=10), nasopharynx (n=2), unknown (n=2), thyroid (n=1) and maxillary sinus (n=1). Since the goal was to create a general HNC RapidPlan model, no differentiation was made regarding primary tumor location. New treatment plans for these patients were created using our in-house developed automatic interactive optimizer (AIO). AIO was developed to produce plans with more consistent OAR sparing than manually optimized plans by automatically adapting the dose-volume objectives throughout the interactive optimization process, ensuring that the same level of attention is given to sparing of each OAR[9]. These plans were created in the Eclipse treatment planning system using the progressive resolution optimizer (PRO) v10.0.28 and anisotropic analytical algorithm (AAA) with a 2.5mm grid. The geometric and dosimetric features of the AIO plans were used to create a RapidPlan model library. This model could then be used to predict achievable OAR DVHs for patients outside the model library, based on their OAR-PTV geometry (Table 1). RapidPlan automates the optimization process by generating a line of optimization objectives just below the inferior boundary of the OAR DVH prediction range.

A standard set of optimization objective priorities, reflecting our institutional practice, was used for all patients. Figure 1 shows the RapidPlan optimization window with various OARs included. The shaded regions represent the OAR DVH prediction ranges, while the dotted lines represent the automatically generated optimization objectives, placed just below the inferior boundary of the DVH prediction range. To prevent underdosing of the PTVs, RapidPlan is designed to place the line objective horizontally in the OAR-PTV overlap volume. In this study, line objectives were generated for the parotid glands, submandibular glands, individual swallowing muscles and oral cavity, while fixed maximum point dose objectives of 37Gy and 39Gy were set for the spinal cord / brainstem and their planning at risk volumes, respectively.

Table 1. Summary of the RapidPlan model characteristics for each individual organ-at-risk (OAR). The model fit R^2 value shows the correlation between the geometric and dosimetric regression parameters. Where relevant, the RapidPlan model configuration algorithm can suggest to include more OARs in the model library. This could for example be in the case where there is large variation in OAR sparing between the plans that are included in the model library. Including more plans could improve the prediction accuracy. Because not every structure was attempted to be spared in the original clinical plan, the number of structures included in the model varies per OAR.

OAR Name	Number of structures included in Model	Suggested Number In Model	Model Fit (R^2)	Regression Model's average $\chi 2$
Salivary Glands				
Contralateral Parotid	90	-	0.77	1.096
Ipsilateral Parotid	86	-	0.52	1.051
Contralateral Submandibular	69	-	0.68	1.094
Oral Cavity	76	-	0.80	1.073
Larynx				
Lower	46	-	0.82	1.125
Upper	26	36	0.76	1.196
Pharyngeal Constrictor Muscles				
Inferior	38	48	0.86	1.229
Medial	24	42	0.58	1.169
Superior	39	48	0.89	1.128
Upper Esophageal Spincter	68	-	0.85	1.147
Cricopharyngeal Muscle	52	-	0.75	1.117

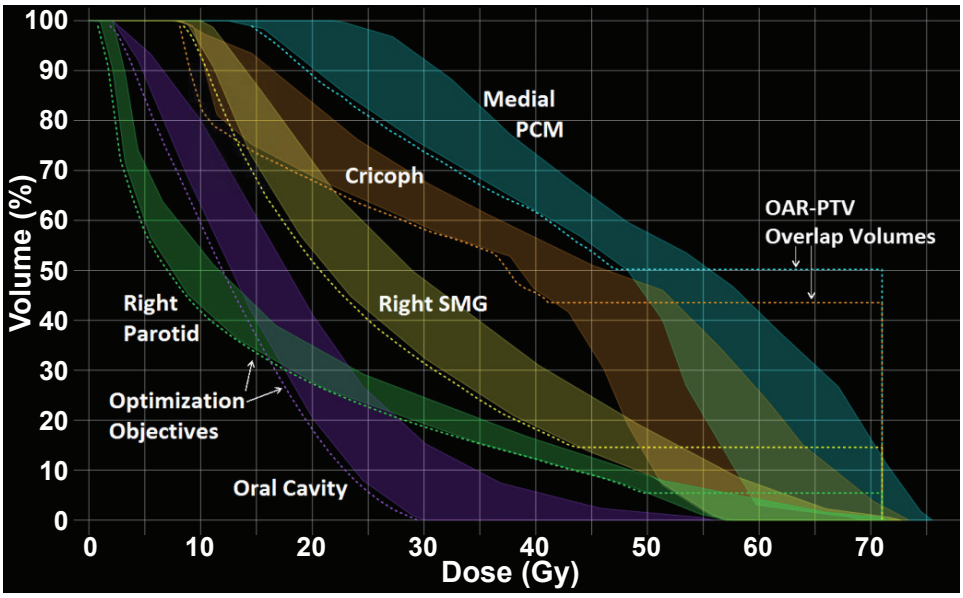


Figure 1. Organ-at-risk (OAR) dose-volume histogram (DVH) prediction ranges (shaded regions) generated by the RapidPlan model and optimization objectives placed along the inferior DVH prediction boundary (dotted lines). To prevent underdosing of the planning target volume (PTV), line objectives are placed horizontally in the portion of the OAR that overlaps with the PTV.

Replanning of dosimetric outliers

The RapidPlan model configuration window provides detailed information regarding geometric and dosimetric outliers that are present in the created model. Although AIO provides an automated and consistent approach to OAR dose reduction during the optimization process, some OARs could still be identified by the model as providing insufficient sparing. This could for example be because a trade-off with sparing of other OARs or strict PTV dose homogeneity criteria prevented better sparing. The consistency of the plans included in the model was therefore improved by two iterations of replanning such dosimetric outliers. In this process, the included OAR DVHs were compared against the range of achievable DVHs predicted by the model. A patient was replanned using RapidPlan if, subjectively judged, a meaningful improvement in sparing was predicted for at least one OAR. The RapidPlan model configuration window assists the user in this process by allowing them to visualize the predicted and achieved DVHs for all OARs included in the model. Cleaning the model in this fashion is intended to improve the consistency of the included plans and lead to more accurate predictions of achievable OAR DVHs for new patients.

As an example, Figures 2a and 3a show two OARs identified as dosimetric outliers because of a discrepancy between the predicted DVH range (shaded blue region) and the achieved DVH (solid blue line), as visualized in the model configuration window. Figures 2b and 3b show the resulting DVH-lines after replanning the corresponding patients using the RapidPlan model, resulting in improved sparing for both OARs and a closer correspondence between the predicted and achieved DVHs. The residual plots (Figures 2c and 3c), indicating the relation between the achieved OAR DVH (y-axis) and the OAR DVH predicted by the RapidPlan model (x-axis), also improved after replanning (Figures 2d and 3d). In total, 19 and 15 plans containing one or more dosimetric outlier OARs were replanned in the first and second iteration, respectively.

The discrepancy between the predicted and achieved DVH-lines (Figures 2a and 3a) was solved by replanning the corresponding patient (indicated by solid blue line) using the RapidPlan model (2b and 3b). The shaded region indicates the DVH prediction range for the selected plan. RapidPlan uses principal component analysis to decompose the shape of the achieved and predicted DVHs in the model library, allowing for a more consistent way to compare the estimated and obtained dosimetry. The residual plot (2c and 3c) shows the correlation between the obtained (DVH principal component score 1) and predicted dosimetry (estimated DVH principal component score 1) after replanning (2d and 3d), indicating that the predicted OAR DVH closely corresponds to the OAR DVH included in the model library. Since more than one upper larynx was identified as an outlier in the two iterations of outlier replanning, more DVHs are noted to change.

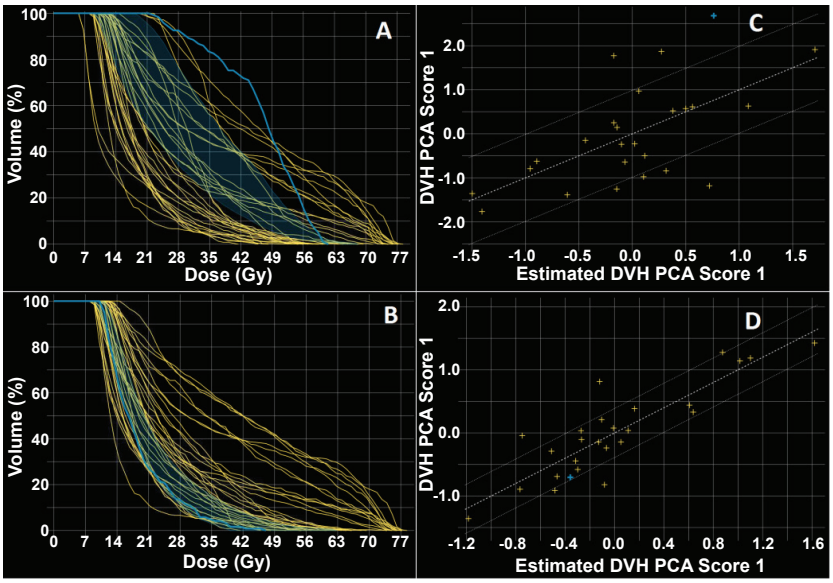


Figure 2. Example of an upper larynx contained in the RapidPlan model that was identified as a dosimetric outlier (Figure A) in the model configuration window. Replanning (B) reduced the discrepancy with the predicted DVH (B). Screenshots taken from the RapidPlan model configuration window.

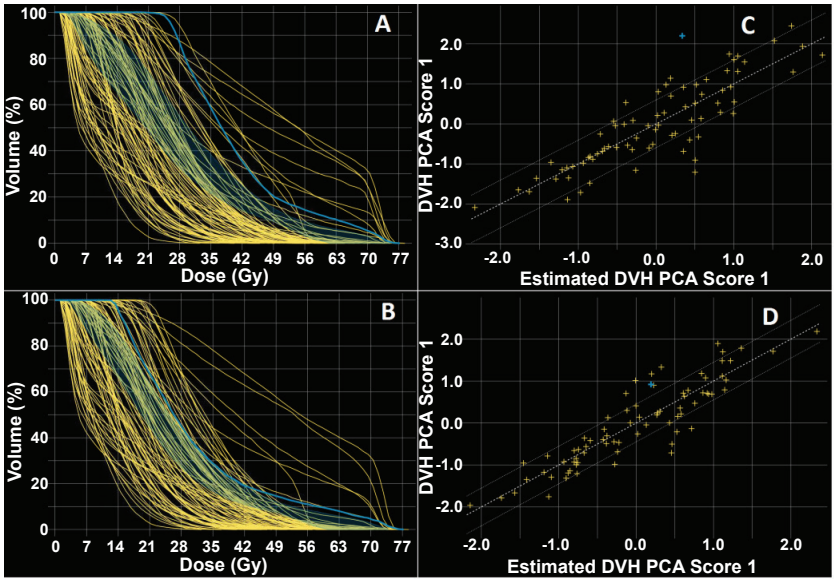


Figure 3. Similar to Figure 2, an oral cavity identified as an outlier in the RapidPlan model library.

Evaluation Group for Plan Quality Checking

The evaluation dataset consisted of the contoured planning CT-scans of 20 HNC patients treated between 2012 and 2013, along with their manual interactively optimized clinical plans (i.e. plans where the planner manually adapted the position of the OAR dose-volume objectives during the optimization process, relative to the position of the DVH-lines). For the evaluation group, primary tumors were oropharynx (n=11), (supra-)glottic larynx (n=5), hypopharynx (n=2), and unknown (n=2). These plans were not part of the model library. The following investigations were performed:

(1) For each OAR spared in the original clinical plan, the mean dose predicted by the RapidPlan model was derived by creating a "mid-prediction" DVH-line running through the middle of the DVH prediction range (Figure 4). To determine the accuracy of the predicted DVH, this mean dose was compared against the mean dose that was obtained when a plan was made using the RapidPlan model.

(2) Because this comparison of mean doses does not reflect possible differences between the mid-prediction DVH-line and the achieved DVH-line, this difference was determined over the entire dose range.

(3) The mid-prediction DVH-line was used to evaluate the quality of AIO plans and manually optimized clinical plans for the 20 patients in the evaluation group. The manually optimized plans were created before AIO was introduced in our clinic, and based on prior experience[7], were generally expected to provide less OAR sparing than the AIO plans. For illustrative purposes, the plans were considered acceptable if the mean dose to the composite (volume weighted) salivary glands, oral cavity and composite swallowing muscles was no more than 3Gy, 5Gy and 7Gy higher, respectively, than the mean dose of the mid-prediction DVH-line. These arbitrary thresholds were chosen to represent clinically meaningful values. A higher threshold value was used for the oral cavity because institutional experience suggest that this structure is subject to more contouring and geometric variability. The threshold was highest for the swallowing muscles because these structures are relatively small and large dose differences are more easily obtained when replanning. Since the OAR DVH prediction by the RapidPlan model takes the geometrical features of the evaluated patient into account, this provides a patient specific approach to plan quality assurance.

Results

Figure 5 shows for the first 3 patients examples of the DVH prediction range estimated by the RapidPlan model along with the DVHs that were achieved when using RapidPlan to create a new treatment plan and the previously created clinical plan DVHs. For the contralateral parotid and submandibular glands, the achieved DVHs were relatively comparable for both planning methods, indicating that the clinical plan spared these structures comparably to the plans that were included in the model. More variation is seen for the upper larynx and oral cavity, where the predicted gains generally were achieved by RapidPlan. The time required to generate the DVH predictions was less than 2 minutes while the optimization time of the RapidPlan plans was typically 10-15 minutes.

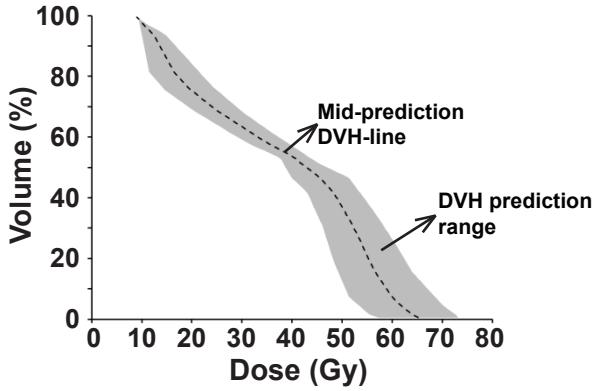


Figure 4. The mid-prediction dose-volume histogram (DVH) line (dashed) running through the middle of the DVH prediction range (shaded region). This was used as a surrogate for the prediction DVH in this study and determined using an in-house developed program coded in Lazarus (*http://www.lazarus.freepascal.org/*).

The maximum dose values to the spinal cord and brainstem were found to be clinically acceptable in all plans. Compared to the manually optimized plans, RapidPlan improved mean composite salivary gland doses by on average 2.0±2.1Gy (range of -2.7Gy to 6.5Gy), mean oral cavity doses by 3.6±3.6Gy (-2.1Gy to 9.7Gy) and mean composite swallowing muscle doses by 5.9±2.9Gy (0.6Gy to 11.4Gy). Figure 6 shows the predicted mean dose plotted against the achieved mean dose for multiple OARs. The solid line represents a linear fit created through all datapoints. The dashed line has a slope of 1 and runs through the origin, meaning that for OARs on this line, the mean dose that was predicted by RapidPlan was exactly achieved. All OARs showed a strong linear correlation between predicted

and achieved mean doses, with R^2 correlation coefficient values ranging from 0.94 for the contralateral parotid gland to 0.99 for the ipsilateral parotid gland. For all OARs combined, the linear fit has a R^2 of 0.97, and a slope greater than 1 (1.08). This indicates that on average, the achieved mean dose was slightly higher than the mean dose that was predicted by RapidPlan. This was likely caused by several OARs with mean doses >40Gy that were located above the dashed fits, indicating that the RapidPlan model overestimated the amount of OAR sparing that could be achieved. These OARs consisted of one ipsilateral parotid gland, one cricopharyngeal muscle and 6 pharyngeal constrictor muscles, overlapping with the PTVs by 43±11%, on average. For these OARs, RapidPlan was unable to accurately predict the amount of OAR sparing that was achieved in the portion of the OAR that overlapped with the elective and boost PTV.

For each OAR, Figure 7a shows the dose difference (ΔDose, y-axis), computed as achieved DVH dose minus mid-prediction DVH dose, plotted against the mid-prediction DVH dose (x-axis). Series of datapoints can be noted running through the graph. These datapoints belong to OARs for which ΔDose changes gradually with dose. More sudden, larger changes in ΔDose, suggesting less accurate predictions, can also be noted, for example in the bottom portion of the graph around 30-40Gy. This graph shows that the prediction is generally more accurate at low OAR doses. At high OAR doses, and therefore at large overlap volumes with the PTVs, RapidPlan often overestimated the amount of sparing that could be achieved (ΔDose values greater than 0). Figure 7b shows the same ΔDose values at 5% OAR volume increments, separated on the basis of the mean OAR dose achieved in the final RapidPlan plan. For all OARs, ΔDose is closer to 0 at higher OAR volumes (typically receiving low doses). At the higher dose regions (OAR volumes < 30%), the amount of achievable sparing is underestimated for OARs with mean doses < 20Gy while it is progressively overestimated for OARs with higher mean doses. This likely resulted from the horizontal placement of the dose-volume objectives at the OAR-PTV overlap volume (Figure 1). Consistent with this hypothesis, the amount of OAR-PTV overlap for the OARs with achieved mean doses of <20Gy, 20-30Gy, 30-40Gy and >40Gy was 1.2±2.0%, 7.7±7.0%, 19.8±9.8% and 35.9±11.9%, respectively.

The OAR DVH predictions generated by the RapidPlan model were used to re-evaluate the quality of the 20 manually optimized clinical plans, benchmarked using the mean dose of the mid-prediction DVH-line. 10 plans were found acceptable and provided sparing of the composite salivary glands, oral cavity and composite swallowing muscles at most 3Gy, 5Gy and 7Gy higher than the mean dose of the mid-prediction DVH-line generated for these structures, respectively. Figure 8 shows the predicted and achieved mean dose values of the

clinical plans that passed (green circles) and failed (red circles) these criteria, combined with the linear fits between the predicted and achieved RapidPlan mean doses shown in Figure 6. The small number of points located below the fit indicate that for the majority of OARs, the RapidPlan model improved OAR sparing over the clinical plan. Using the same thresholds, 4 AIO plans did not pass the QA, all because of inferior oral cavity sparing. The salivary glands and swallowing muscles were within the chosen thresholds for OAR sparing in all AIO plans. As expected from the high R^2 values, after replanning of the patients that violated the criteria using RapidPlan, all resulting OAR mean doses fell within the respective threshold range from the predicted mean doses.

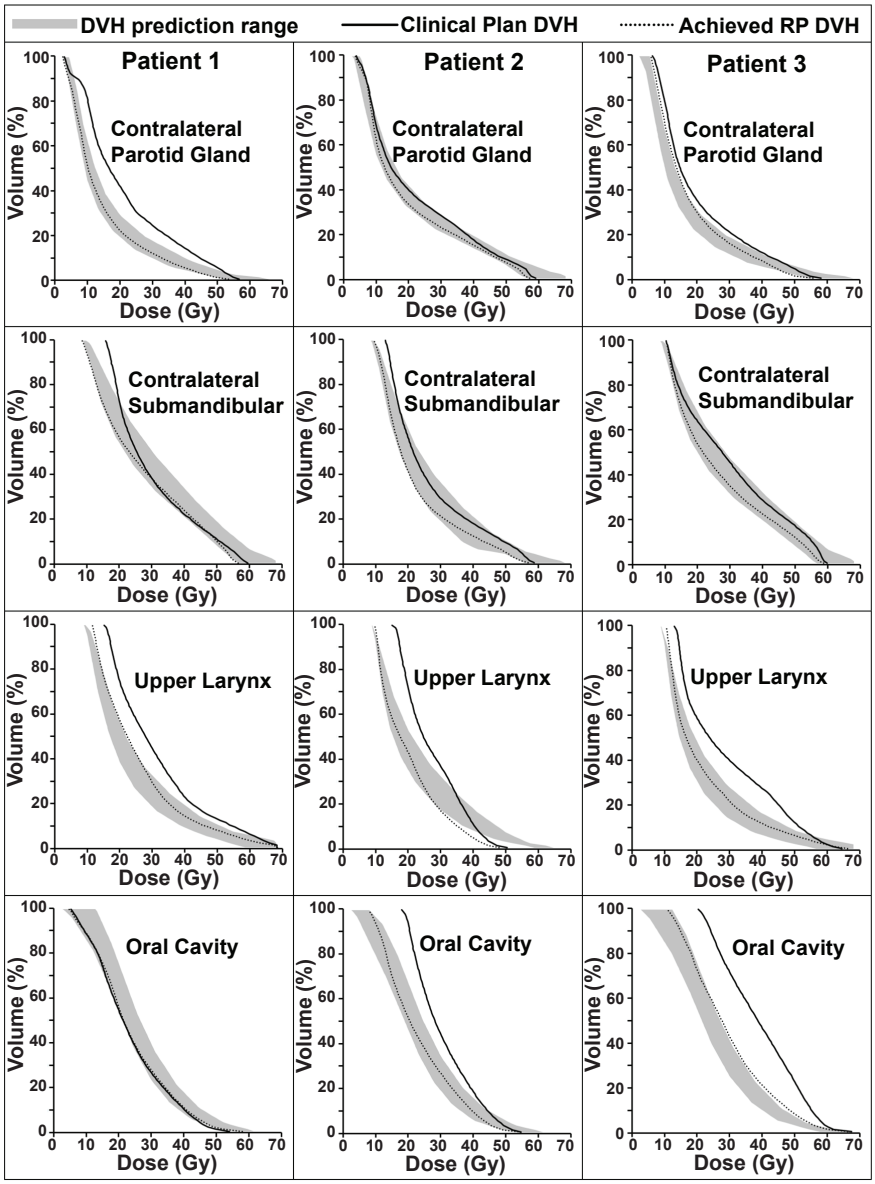


Figure 5. Examples of predicted dose-volume histogram (DVH)-ranges (shaded regions) and achieved DVH-lines for multiple organs-at-risk (OARs) of three patients. The solid lines represent the DVHs that were achieved in the previously created clinical plans, while the dotted lines indicate the DVHs that were obtained when the RapidPlan model was used to create a new treatment plan. The mid-prediction DVH-line used for the analysis in the present report is located in the middle of the shaded region.

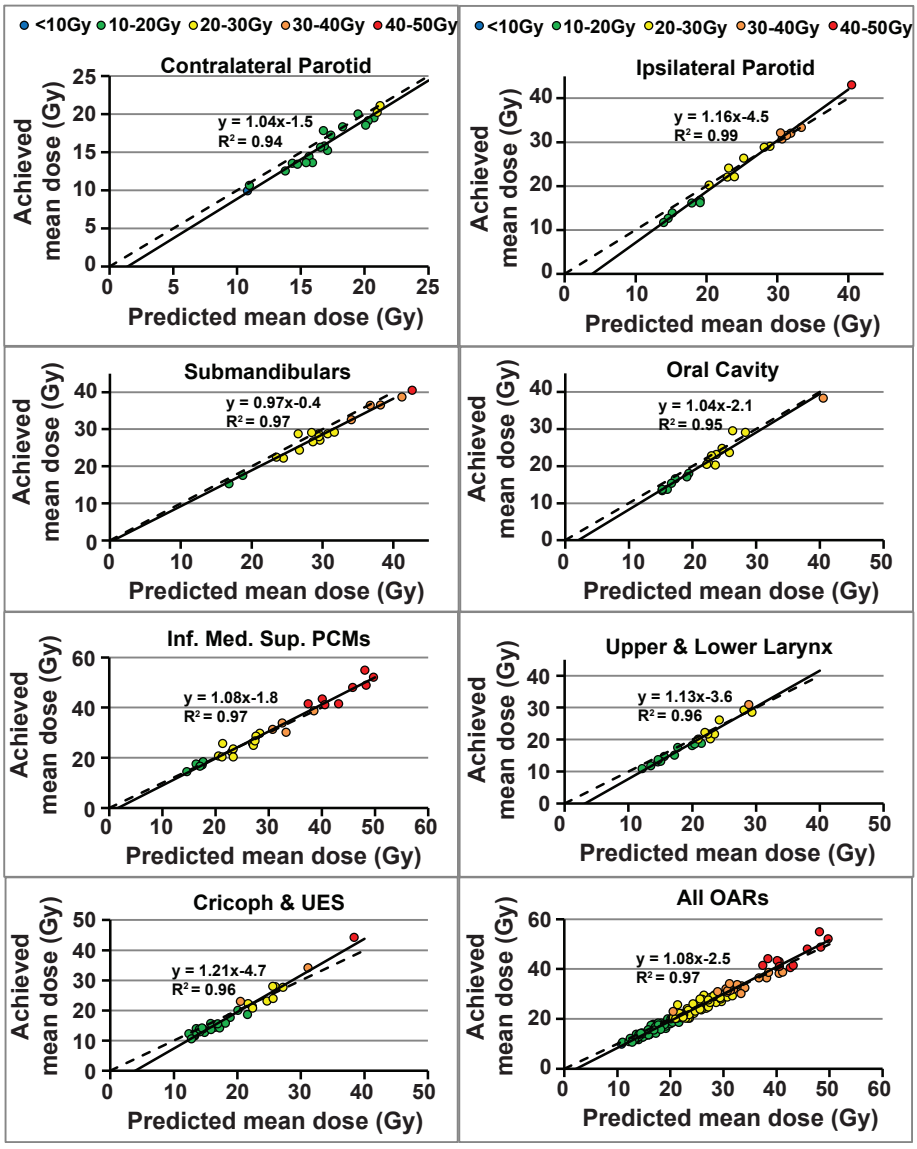


Figure 6. For multiple organs-at-risk (OARs), the correlation between predicted (x-axis) and achieved (y-axis) mean OAR doses. The solid lines represent fits created through all datapoints, while the dashed line indicates a linear fit through the origin. The R^2 values indicates the goodness-of-fit of the solid line with the datapoints. For conciseness, some individual swallowing muscles and the contralateral, and ipsilateral submandibular gland are analyzed together in these graphs. The number of OARs included in these graphs can vary depending on whether they were designated to be spared in the original clinical plan.

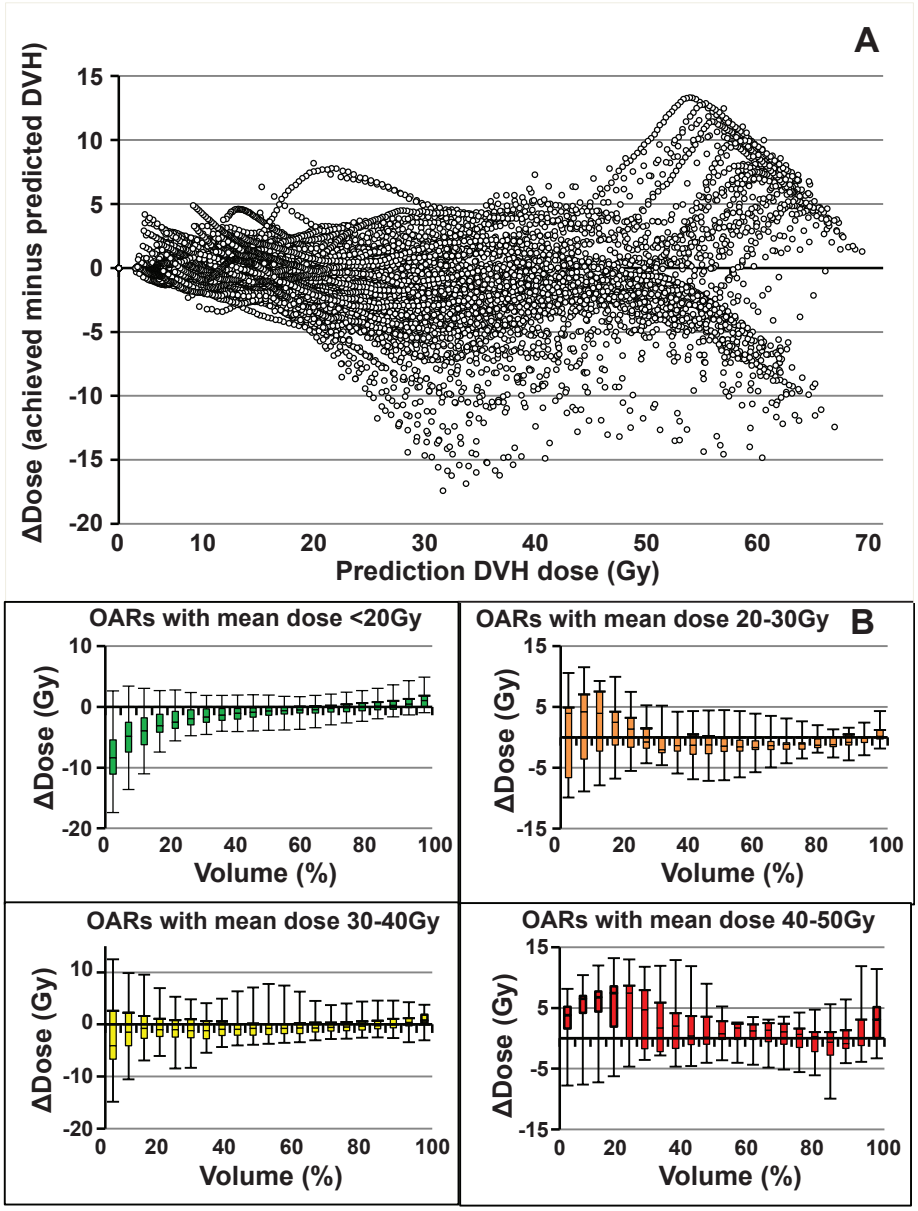


Figure 7. The prediction accuracy of RapidPlan along the dose-volume histogram (DVH)-line. A) Dose difference between the achieved and predicted DVH-line (ΔDose, y-axis) plotted against the dose of the predicted DVH-line (x-axis). B) Box-whisker plots of ΔDose as a function of organ-at-risk (OAR) volume (x-axis) for four different ranges of OAR mean doses. Lower OAR volumes are typically associated with high doses.

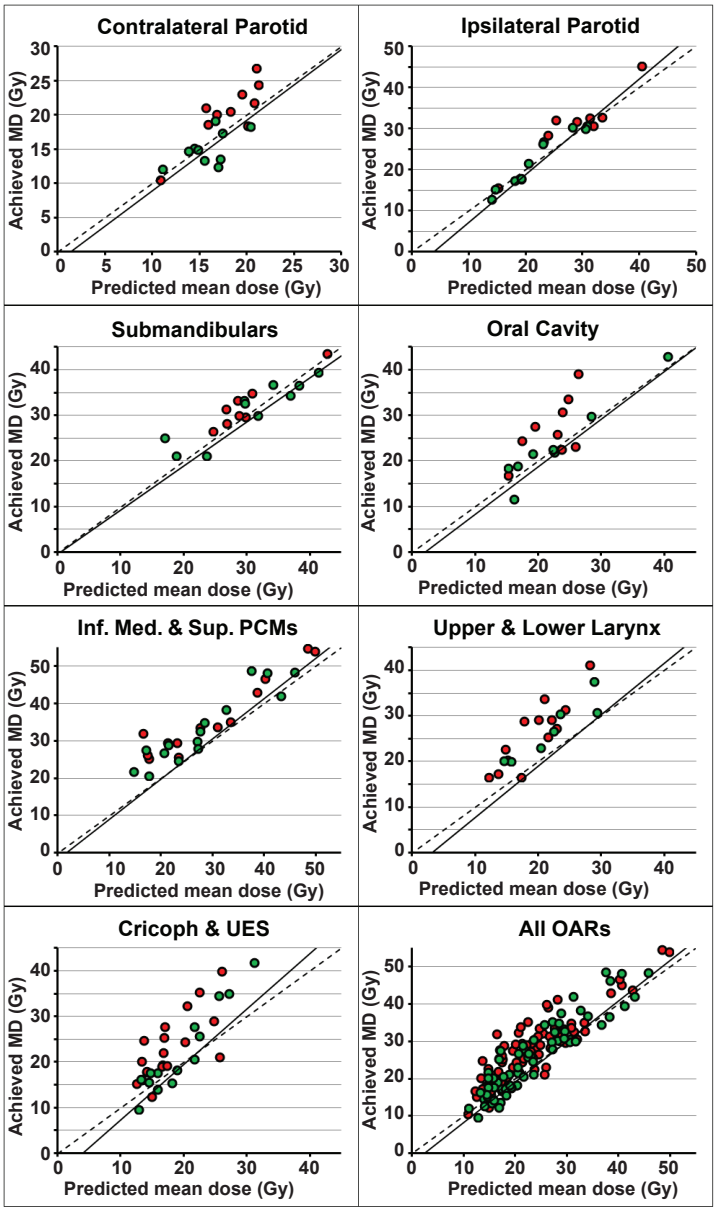


Figure 8. The organs-at-risk (OARs) of the clinical plans that passed (green circles) and failed (red circles) the evaluation criteria, along with the linear fit between the RapidPlan predicted and achieved mean dose found in Figure 6. Thresholds of 3Gy, 5Gy and 7Gy were used for quality evaluation of the composite (volume-weighted) salivary glands, oral cavity and composite swallowing muscles, respectively. Datapoints of individual OARs with similar predicted and achieved mean doses could still fail the criteria because the analysis was done based on composite OARs.

Discussion

This study assessed the potential to use OAR DVH predictions generated by a RapidPlan model to evaluate the dosimetric quality of plans that were not included in the model. For all OARs, strong correlations were found between mean OAR doses predicted by the model and mean doses achieved after the model was used to guide the creation of a new treatment plan, with linear fit slopes close to one. This means that in general, achievable OAR mean doses could be determined in advance by using the model solely to predict DVHs, without requiring the creation of an actual treatment plan. Although previous investigations showed the potential of RapidPlan to improve plan quality[5–7], the present study is the first to evaluate the accuracy of the dosimetric predictions. This is relevant because improved plan quality does not necessarily imply that accurate predictions were generated, whilst accurately modeling achievable OAR doses is an important prerequisite for using RapidPlan as a plan QA tool. The prediction accuracy was determined for mean OAR doses because such doses have been shown to correlate to late toxicity for HNC patients[22–26].

Since the created RapidPlan model was found able to accurately predict the achievable mean OAR doses (Figure 6), it could be used to evaluate the quality of previously created (manually optimized) clinical plans. Under the test conditions, this suggested that 50% of the evaluated plans contained a composite OAR for which the sparing could be improved. These predicted gains in plan quality were not unexpected since the RapidPlan model was made using an AIO plan library, and AIO plans have been previously shown to provide improved OAR sparing over manual interactively optimized plans, along with being optimized in a more consistent fashion[9]. Consistent with this, only 4 out of 20 AIO plans were rejected because oral cavity sparing could have been improved, while salivary gland and swallowing muscle sparing was acceptable in all plans. The predicted improvements in plan quality were validated by replanning the patient using the RapidPlan model. The agreement was consistent with the high R^2 values between the predicted and achieved mean OAR doses, and thus demonstrated successful application of RapidPlan-predicted DVH metrics as a QA tool. It is important to note that the high-level correlation between predicted and achieved dosimetry also allows the RapidPlan model to be used to evaluate the plans included in the model library. Replanning sub-optimal plans would allow for the continuous and consistent improvement of the model library and the resulting plan quality.

Although on average, RapidPlan could accurately predict mean doses, predictions could deviate in the high-dose regions for OARs with both low (<20Gy) and high (>40Gy) mean doses (Figure 7). For the latter, RapidPlan overestimated the amount of achievable OAR

sparing, likely because RapidPlan placed the line objective horizontally in the region of OAR-PTV overlap. In plans with multiple dose levels, OARs that only overlap with the lower dose PTV get no optimization objectives to restrict doses above the prescribed dose for this PTV (Figure 1). Although this may not be as important in organs where toxicity is correlated with mean OAR dose, it could be an issue when high dose volumes are considered predictors of toxicity. If in future releases of RapidPlan, line objectives are also modeled for the part of the OAR that overlaps with the PTV, correlations between predicted and achievable OAR DVHs should improve.

Clinical trials often include treatment plans submitted by a large number of institutes. The difficulty of determining whether sufficient OAR sparing is reached for individual patients may lead to the inclusion of poor treatment plans, even though the magnitude of achieved OAR sparing may influence the outcome of the study[11,27]. RapidPlan models that consist of a wide range of patient plans in which a consistently good level of OAR sparing was reached allow for fast, patient-specific evaluation of OAR DVHs for plans submitted to the study. Figure 9 suggests a workflow to quickly decide whether a submitted plan provides sufficient sparing based on its OAR-PTV geometry. In addition, the present results show the feasibility for clinical trials to supply participating centers with a RapidPlan model to ensure that all created plans are according to the specified guidelines. It is important to note that the validity of a RapidPlan model needs to be re-evaluated when dose prescriptions or optimization priorities change over time. Trial-specific model libraries may be required depending on the planning criteria.

Accurate DVH predictions made by RapidPlan models could similarly be used to benchmark the quality of plans that were created using alternative delivery techniques, such as proton therapy[28], without requiring the creation of photon plans. A comparable study was recently performed that evaluated the potential of an IMRT model from one institute that was made using one IMRT technique, to aid IMRT planning of another institute and predict achievable DVHs with another IMRT technique[18]. They found that a fixed gantry model could accurately predict the median dose of the parotid gland in Tomotherapy plans, which indicated that it was spared similarly by both institutions. In addition, predictions of median dose reductions to some OARs in the Tomotherapy plans could be achieved after replanning.

Potential limitations of the current study include the following. RapidPlan indicated that for some swallowing muscles, more training cases containing these structures were required to improve the model (Table 1). This may have influenced the predicted and obtained mean

doses. All plans included in the model were created using the same number of arcs, and similar collimator angles and field sizes for most plans. If plans created using different field settings are evaluated, RapidPlan may predict OAR doses that are only realizable after changing the field settings. Additionally, different optimization and dose calculation algorithms were used to create the plans contained in the RapidPlan model library (PRO and AAA v10.0.28) and the RapidPlan plans (photon optimizer [PO] and AAA v13.5.33). Small improvements in OAR sparing when using newer algorithms were found previously[7]. Part of the gains achieved by RapidPlan may therefore be attributed to the use of the newer optimization and dose calculation algorithms, and in effect, the RapidPlan model predictions slightly underestimated the gains that could be achieved. Although our study showed good correlations between achieved and predicted OAR doses, determining the optimality of the plans included in the model library or of the resulting plans was beyond the scope of the present study. Future studies could incorporate retrospective RapidPlan QA of completed clinical trials. If a RapidPlan model were to be used as a planning QA tool for clinical trials, time and effort would need to be invested to ensure appropriate model composition and quality. Finally, if changes in treatment planning techniques lead to improvements in plan quality, these improved plans would need to be incorporated into updated RapidPlan libraries to ensure that the QA model remained state of the art and of sufficiently high quality.

The present study showed that RapidPlan models can accurately predict achievable mean doses for most OARs, enabling them to be used to benchmark the quality of existing plans. However, the predictions may be less accurate for OARs that overlap substantially with the PTVs and may not accurately reflect the shape of the DVH-line. Addressing these limitations should improve the prediction accuracy and the potential of RapidPlan for plan quality assessment.

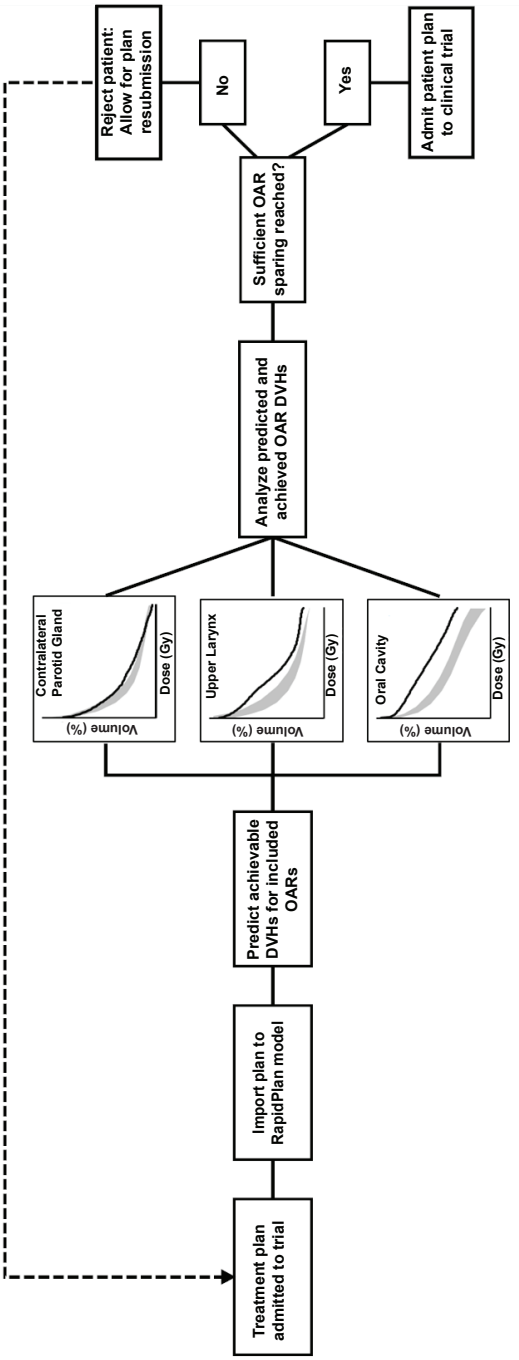


Figure 9. A flowchart proposing how the organ-at-risk (OAR) dose-volume histogram (DVH) predictions generated by a RapidPlan model could be used for fast plan quality assurance (QA) for clinical trials.

References

1 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

2 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

3 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

4 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

5 A. Fogliata, P.M. Wang, F. Belosi, A. Clivio, G. Nicolini, E. Vanetti, and L. Cozzi, "Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer.," Radiat. Oncol. **9**(1), 236 (2014).

6 A. Fogliata, F. Belosi, A. Clivio, P. Navarria, G. Nicolini, M. Scorsetti, E. Vanetti, and L. Cozzi, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer.," Radiother. Oncol. **113**(3), 385–91 (2014).

7 J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Evaluation of a knowledge-based planning solution for head and neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **91**(3), 612–20 (2015).

8 B. Wu, T. McNutt, M. Zahurak, P. Simari, D. Pang, R. Taylor, and G. Sanguineti, "Fully automated simultaneous integrated boosted-intensity modulated radiation therapy treatment planning is feasible for head-and-neck cancer: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **84**(5), e647–53 (2012).

9 J.P. Tol, M. Dahele, J. Peltola, J. Nord, B.J. Slotman, and W.F.A.R. Verbakel, "Automatic interactive optimization for volumetric modulated arc therapy planning," Radiat. Oncol. **10**(1), 75 (2015).

10 E.A. Miles, C.H. Clark, M.T.G. Urbano, M. Bidmead, D.P. Dearnaley, K.J. Harrington, R. A'Hern, and C.M. Nutting, "The impact of introducing intensity modulated radiotherapy into routine clinical practice.," Radiother. Oncol. **77**(3), 241–6 (2005).

11 L.J. Peters, B. O'Sullivan, J. Giralt, T.J. Fitzgerald, A. Trotti, J. Bernier, J. Bourhis, K. Yuen, R. Fisher, and D. Rischin, "Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02.," J. Clin. Oncol. **28**(18), 2996–3001 (2010).

12 J. Wang, W. Chen, M. Studenski, Y. Cui, A.J. Lee, and Y. Xiao, "A semi-automated tool for treatment plan-quality evaluation and clinical trial quality assurance.," Phys. Med. Biol. **58**(13), N181–7 (2013).

13 K.L. Moore, R. Schmidt, V. Moiseenko, L.A. Olsen, J. Tan, Y. Xiao, J. Galvin, S. Pugh, M.J. Seider, A.P. Dicker, W. Bosch, J. Michalski, and S. Mutic, "Quantifying Unnecessary Normal Tissue Complication Risks due to Suboptimal Planning: A Secondary Study of RTOG 0126.," Int. J. Radiat. Oncol. Biol. Phys. **92**(2), 228–35 (2015).

14 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

15 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

16 X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

17 K.L. Moore, R.S. Brame, D.A. Low, and S. Mutic, "Experience-based quality control of clinical intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **81**(2), 545–51 (2011).

18 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

19 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

20 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

21 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

22 I. Beetz, C. Schilstra, A. van der Schaaf, E.R. van den Heuvel, P. Doornaert, P. van Luijk, A. Vissink, B.F.A.M. van der Laan, C.R. Leemans, H.P. Bijl, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, and J.A. Langendijk, "NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors.," Radiother. Oncol. **105**(1), 101–6 (2012).

23 T. Dijkema, C.P.J. Raaijmakers, R.K. ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M. a Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

24 J.O. Deasy, V. Moiseenko, L. Marks, K.S.C. Chao, J. Nam, and A. Eisbruch, "Radiotherapy dose-volume effects on salivary gland function.," Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl), S58–63 (2010).

25 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F.A.M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

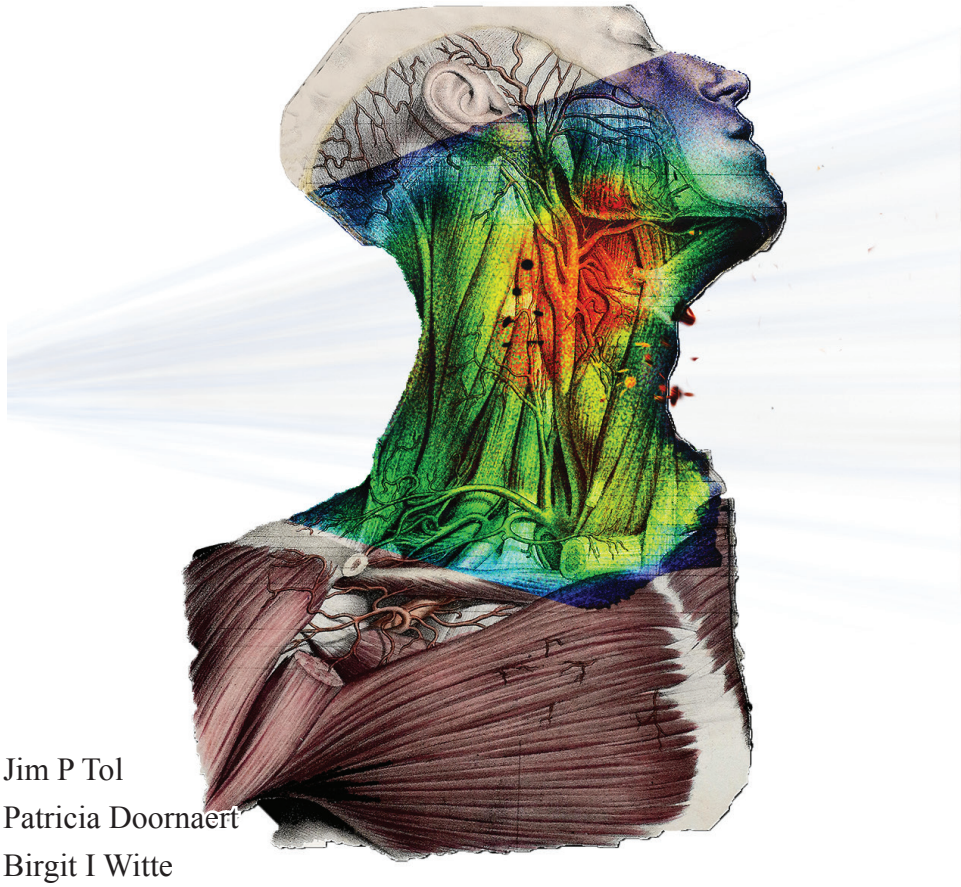
26 C.A. Murdoch-Kinch, H.M. Kim, K.A. Vineberg, J.A. Ship, and A. Eisbruch, "Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **72**(2), 373–82 (2008).

27 D.C. Weber, C.W. Hurkmans, C. Melidis, W. Budach, J.H. Langendijk, L.J. Peters, V. Grégoire, P. Maingon, and C. Combescure, "Outcome impact and cost-effectiveness of quality assurance for radiotherapy planned for the EORTC 22071-24071 prospective study for head and neck cancer.," Radiother. Oncol. **111**(3), 393–9 (2014).

28 D.L.J. Barten, J.P. Tol, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Comparison of organ-at-risk sparing and plan robustness for spot-scanning proton therapy and volumetric modulated arc photon therapy in head-and-neck cancer," Med. Phys. **42**(11), 6589–6598 (2015).

Chapter 10

A longitudinal evaluation of improvements in radiotherapy treatment plan quality for head and neck cancer patients



Jim P Tol
Patricia Doornaert
Birgit I Witte
Max Dahele
Ben J Slotman
Wilko FAR Verbakel

Radiotherapy and Oncology **119**(2), 337-343 (2016).

Abstract

Purpose

To investigate changes in head and neck cancer (HNC) plan quality following the introduction of new technologies and planning techniques in the last decade.

Materials and Methods

Thirty plans were selected from each of four successive periods (P). P1: 7-field static intensity modulated radiotherapy (IMRT) with parotid gland sparing; P2: dual-arc volumetric modulated arc therapy (VMAT, also used in P3-P4), including submandibular gland sparing; P3: inclusion of individual swallowing muscles and attempts to further reduce parotid and oral cavity doses through manual interactive optimization; P4: containing the same organs-at-risk (OARs) as P3, but automatically interactively optimized. Plan benchmarking included mean salivary gland / swallowing muscle / oral cavity (D_{sal} / D_{swal} / D_{oc}) doses. Differences in mean doses between the periods were analyzed by an ANCOVA, taking geometric differences across the periods into account.

Results

Compared to P1, P2 plans improved D_{sal} by 3.4Gy on average. P3 improved D_{sal} / D_{swal} / D_{oc} by 6.9Gy / 11.5Gy / 7.2Gy over P2, showing that D_{swal} and D_{sal} could be improved simultaneously. In P4, D_{oc} / D_{swal} slightly improved over P3 by 1.7Gy / 3.8Gy. Improved OAR sparing in P3 / P4 did not come at the cost of increased dose deposition elsewhere and planning target volume (PTV) dose homogeneity was similar.

Conclusions

New technologies and planning techniques were successfully implemented into routine clinical care and improved HNC plan quality.

Introduction

Radiotherapy treatment planning for locally advanced head and neck cancer (HNC) evolved over the years from only attempting spinal cord sparing[1], to sparing multiple additional organs-at-risk (OARs), including the parotid and submandibular glands, individual swallowing muscles and oral cavity[2]. Minimizing salivary gland and oral cavity doses is important for reducing the severity and incidence of xerostomia and oral mucositis, respectively[3–5], while recent investigations highlight the importance of swallowing muscle sparing to prevent dysphagia[2,6–8]. Increased treatment plan complexity has been facilitated by advances in dose delivery techniques such as intensity modulated radiotherapy (IMRT) and volumetric modulated arc therapy (VMAT) which allow for the shaping of more complicated dose distributions[9]. However, IMRT and VMAT plan quality can vary substantially between planners and radiotherapy centers[10–13], partly due to the learning curve for advanced planning and partly because the achievable level of OAR sparing is unknown before starting the planning process. Automated solutions to generating consistent, high-quality treatment plans are therefore attracting interest[14–19].

Treatment planning improvements are often based on traditional planning studies, comparing plans created by different delivery or planning techniques on the same patient cohort, frequently leading to the adoption of the better technique for future patients. However, once new techniques are introduced in routine clinical practice, the realized gains in plan quality and dosimetry are rarely investigated.

Over the last decade, treatment planning for HNC at the VU University Medical Center evolved from static gantry IMRT plans solely attempting parotid gland, spinal cord and brainstem sparing, to automatically optimized dual-arc VMAT plans sparing the parotid and submandibular glands, up to seven individual swallowing muscles, and the oral cavity. However, moving from controlled planning studies to routine clinical practice creates the risk that increasing plan complexity might lead to unexpected negative consequences. For example, sparing new OARs might degrade sparing of previously included OARs, dose deposition in normal tissue might increase, or dose conformity might be compromised. To investigate the impact of the successive introduction of new technologies and planning techniques, combined with the sparing of more OARs, on routine plan quality, we performed an analysis of longitudinal dosimetric trends by comparing four periods from 2005-2015. Endpoints included sparing of new and prior OARs, dose conformity, planning target volume (PTV) dose coverage and homogeneity, and dose deposition in the body.

Materials and Methods

Patient selection

Treatment plans of 120 locally advanced HNC patients were selected from four different time periods (30 patients per period). Each period is characterized by the introduction of a new technology or planning approach. Geometric variability was minimized by only including patients with primary tumors located in the tonsillar region or lateral pharyngeal wall of the oropharynx.

Period 1 (P1, May 2005 to October 2008) patients were typically treated using seven equidistant coplanar IMRT fields with emphasis on PTV dose coverage and parotid gland, spinal cord and brainstem sparing.

P2 plans (July 2008 to May 2010) cover the introduction of VMAT (RapidArc™, Eclipse™ treatment planning system, Varian Medical Systems, Palo Alto, USA) for HNC patients in July 2008[20,21]. Contralateral submandibular gland sparing was introduced in 2009[22].

In P3 (August 2012 to October 2013) individual swallowing muscles were included for sparing[23] and the spinal cord maximum dose constraint was placed below 40Gy. Further dose reductions to the salivary glands were also attempted, facilitated by additional instructions and planner training in performing interactive optimization, and improved planning protocols. Additionally, while the oral cavity was occasionally included in P1-P2, routine sparing commenced halfway through P3. A "continue previous optimization" (CPO) was introduced after the first RapidArc optimization and subsequent dose calculation. The CPO is designed to compensate for differences between the fast dose calculation algorithm used during optimization, and the final, high resolution dose calculation, and generally improves PTV dose homogeneity[24].

In P1-P3, manual interactive optimization was performed using 3-5 optimization objectives for each parallel OAR, attempting to maintain these a few Gy from the dose-volume histogram (DVH)-lines that are displayed throughout the optimization process. This ensures that dose reductions to respective OARs are always attempted during optimization. Maximum dose constraints were used and not changed for the spinal cord and brainstem.

P4 (February 2014 to May 2015) included VMAT plans optimized using the in-house developed automatic interactive optimizer (AIO)[19]. The large number of structures added to the optimization in P3 increased the difficulty of manual interactive optimization. AIO automates the optimization process by automatically positioning the OAR objectives at a ~3Gy distance from their respective DVH-lines, reducing the potential for user variability,

leading to more consistent sparing of all included OARs[19]. Since February 2014, AIO has been used clinically to optimize all HNC treatment plans.

Before 2007, a no-action level image-guidance protocol based on MV-imaging was standard[25]. Between 2007 and April 2014, this was replaced by a no-action level protocol based on orthogonal kV-imaging, with daily on-line orthogonal kV-imaging being used increasingly. After April 2014, daily on-line orthogonal kV-imaging became routine. A similar thermoplastic immobilization shell was used in all periods. PTV margins were 3mm / 4mm in P1-P2 / P3-P4. Additionally, from 2010 onwards (including the last 3 patients from P2), PTV margins were increased to 5mm below the clavicles to account for greater caudal positioning uncertainty (positioning of the cranially located boost PTV [PTV_B] took priority).

The optimization and dose calculation algorithms also changed over time. The first 24 plans in P1 used the Helios and pencil-beam convolution algorithm (PBCA) v7.2.34 optimization and dose calculation algorithms, respectively, while the remaining plans used the dose-volume optimizer (DVO) and PBCA v8.1.14. P2 plans used the progressive resolution optimizer (PRO) v8.2.23-v8.6.15, and AAA v8.2.23 to v8.6.15. PRO and AAA v10.0.28 were used for all patients in P3. Finally, PRO v10.0.28, and AAA v10.0.28-Acuros®XB v11.0.31 were used in P4.

All RapidArc plans (P2-P4) were created using two coplanar arcs. In patients receiving treatment prior to 2012, not all OARs were contoured. For this analysis, missing OARs were delineated by a HNC radiation oncologist to retrospectively evaluate the dose in all OARs regardless of whether these structures were spared during treatment. Since the oral cavity was subject to considerable contouring variation, this structure was re-contoured for all patients. Salivary glands included the contralateral (CL) / ipsilateral (IL) parotid and submandibular glands (SMGs). Swallowing muscles were delineated following Christianen et al.[23]. To simplify dose reporting, composite structures were created consisting of all salivary glands ($comp_{sal}$) and swallowing muscles ($comp_{swal}$). For optimization and dose reporting, the PTV was either cropped 6mm beneath the skin (P1 and one patient in P2), or a 5mm 'virtual' bolus region[26] was used.

Dose prescription

In all plans, 70Gy was prescribed to PTV_B, delivered as a simultaneous integrated boost in 35 fractions of 2Gy. PTV_B consisted of the gross tumor volume and biopsy proven positive lymph nodes with a 5mm margin for microscopic disease (edited for anatomical boundaries) and a 3-5mm PTV margin. Elective PTV (PTV_E) prescribed dose was 54.25Gy

for 78% of patients, while 57.75Gy was prescribed for 7 / 19 patients in P1 / P2. PTV_E consisted of elective nodal regions with a 3-5mm margin, minus a 5mm transition region (PTV_T), created to allow for the necessary dose fall-off between PTV_B and PTV_E. Tumor and affected lymph node contouring used all available diagnostic information; MRI and PET-CT scans were (rigidly) co-registered. Table 1 summarizes the patient characteristics, while Table 2 summarizes the OAR, PTV and OAR-PTV overlap volumes in the different periods. Since PTV_B was located centrally in a number of patients, this resulted in mean doses >60Gy to both SMGs. In these cases, both SMGs were considered to be ipsilateral, leading to a fairer comparison of SMG doses.

Study endpoints

Plans from the periods were compared using the following metrics, averaged over the 30 plans of each period; (i) Mean dose to individual OARs and oral cavity / composite salivary glands / swallowing muscles (D_{oc} / D_{sal} / D_{swal}), (ii) PTV_B / PTV_E volumes receiving 95% / 107% of prescribed dose (V95% / V107%), along with their homogeneity indices (HI = 100% × [D2% − D98%] / D50%) and conformity indices (CI) defined as the isodose volumes receiving prescribed PTV doses divided by PTV volumes. (iii) Dose deposition in normal tissue was analyzed using V5Gy / V30Gy / V50Gy and mean dose of the Body – PTV volume (cropped 2cm above and below the PTVs). Additionally, the mean dose to 1cm ring structures, created 0cm, 1cm and 2cm around PTV_B and PTV_E, excluding delineated OARs, was determined.

Statistical Analysis

The data is summarized per period using means ± standard deviations and confidence intervals. Differences were assessed via an ANOVA, while pairwise differences were investigated using independent sample t-tests. To correct differences in OAR dose for variability in OAR, PTV and OAR-PTV overlap volumes, and prescribed PTV_E dose between the periods, an ANCOVA was performed. In case of skewed data, the Kruskal-Wallis test was used. All statistical analyses were performed in SPSS v22 (IBM Corp., Armonk, NY, USA), with significance set at $p < 0.05$.

Table 1. For all considered periods, the approach to planning, tumor stages and fractionation. Thirty patients were included per period and all included patients received bilateral irradiation.

Period	1	2	3	4
	Parotid gland sparing, 6/7-field IMRT	Submandibular gland sparing, dual arc VMAT	Swallowing muscle sparing, dual arc VMAT	Automatically optimized, dual arc VMAT
	Tumor Staging			
T4	12 (N0 = 4 / N1 = 3 / N2 = 4)	12 (N0 = 2 / N1 = 3 / N2 = 5 / N3 = 1)	3 (N0 = 2 / N1 = 1)	8 (N0 = 1 / N1 = 4 / N2 = 3)
T3	10 (N0 = 2 / N1 = 2 / N2 = 6)	6 (N0 = 1 / N1 = 1 / N2 = 4)	3 (N2 = 2 / N1 = 1)	10 (N0 = 2 / N1 = 6 / N2 = 2)
T2	7 (N0 = 1 / N1 = 2 / N2 = 4)	11 (N0 = 3 / N1 = 2 / N2 = 6)	19 (N0 = 1 / N1 = 3 / N2 = 15)	9 (N0 = 3 / N1 = 2 / N2 = 4)
T1	1 (N1 = 1)	1 (N0 = 1)	4 (N1 = 1 / N2 = 3)	2 (N2 = 2)
Tx	-	-	1 (N2 = 1)	1 (N2 = 1)
	Prescription doses / Fractionation			
Boost PTV	70Gy / 35	70Gy / 35x2Gy	70Gy / 35x2Gy	70Gy / 35x2Gy
Elective PTV	54.25Gy / 35 (n=23) 57.75Gy / 35 (n=7)	54.25Gy / 35 (n=11) 57.75Gy / 35 (n=19)	54.25Gy / 35 (n=30)	54.25Gy / 35 (n=30)

Table 2. For all considered periods, the size of planning target volumes (PTVs) and organs-at-risk (OARs), and the OAR-PTV overlap volumes, averaged over all 30 patients. Data is presented by means ± standard deviations (range).

Period	1	2	3	4
	PTV Volumes (cm^3)			
Boost	220.5 ± 102.1 (68.2 to 513.1)	233.7 ± 167.7 (39.8 to 801.8)	184.3 ± 55.2 (94.5 to 315.0)	192.9 ± 104.1 (36.8 to 420.8)
Elective	368.2 ± 120.0 (160.6 to 668.1)	370.0 ± 105.9 (229.9 to 676.4)	387.7 ± 85.1 (238.9 to 592.6)	333.5 ± 58.3 (240.1 to 484.1) #
Transition	36.5 ± 20.6 (6.8 to 101.0)	44.5 ± 35.5 (3.7 to 169.5)	75.5 ± 27.0 (22.8 to 144.3) † *	70.1 ± 37.2 (18.6 to 167.3) † *
Elective + Transition	404.7 ± 133.1 (194.9 to 743.6)	414.5 ± 118.5 (268.3 to 719.1)	463.2 ± 98.0 (284.0 to 697.2)	403.7 ± 60.8 (282.0 to 571.2) #
Combined PTV	625.2 ± 221.0 (263.1 to 1237.8)	648.2 ± 212.8 (372.6 to 1143.5)	647.4 ± 129.2 (378.5 to 878.5)	596.6 ± 133.0 (336 to 882.6)
	OAR Volumes (cm^3)			
Composite Salivary Glands	76.2 ± 20.8 (49.3 to 123.5)	69.4 ± 18.7 (36.2 to 112.8)	82.3 ± 25.3 (31.4 to 130.5) †	75.0 ± 16.1 (38.9 to 107.6) †
Composite Swallowing Muscles	31.1 ± 7.9 (18.5 to 50.8)	31.7 ± 9.1 (17.8 to 51.6)	37.4 ± 9.5 (21.1 to 54.6) † *	34.9 ± 11.1 (23.1 to 71.7) † *
Oral Cavity	122.3 ± 47.0 (43.5 to 270.3)	92.4 ± 29.1 (39.8 to 150.1) *	102.5 ± 25.0 (21.2 to 137.2) *	104.7 ± 51.5 (21.1 to 261.3) *
	OAR-PTV overlap volumes (%)			
Composite Salivary Glands	28.6 ± 11.5 (8.4 to 52.4)	27.4 ± 13.2 (6.5 to 62.5)	23.9 ± 8.6 (9.3 to 42.5)	22.4 ± 9.8 (7.7 to 50.0)†
Composite Swallowing Muscles	26.9 ± 20.6 (7.6 to 86.2)	32.8 ± 23.4 (8.5 to 91.6)	21.1 ± 11.5 (5.1 to 50.1)†	21.9 ± 13.6 (0.7 to 61.5)†

*, † and # indicate statistically significant differences ($p<0.05$) with respect to period 1, period 2 and period 3, respectively. Evaluated using independent sample t-tests.

Results

Maximum serial OAR doses were clinically acceptable for all patients, although average decreases of 8.7Gy / 12.2Gy were noted from P1-P3 for the spinal cord / brainstem, due to the use of lower dose constraints from P3 onwards. Figure 1 shows gradual improvements of D_{sal}, D_{swal} and D_{OC} throughout the periods. Averaged DVHs for a number of structures are shown in Figure 2.

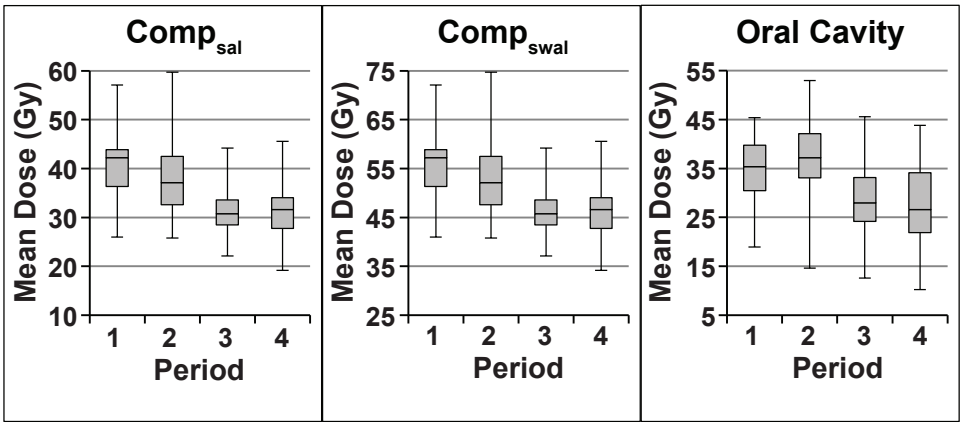


Figure 1. Box-whisker plots showing the spread of mean dose values for composite salivary glands ($comp_{sal}$), swallowing muscles ($comp_{swal}$) and oral cavity.

Table 3 summarizes averaged mean dose differences in P1-P3 with respect to P4. Correction was necessary for 10 out of 14 OARs to account for differences between the selected patients (see Table footnote). The corrections substantially influenced the mean doses for some OARs. Correcting the CL-SMGs for higher PTV overlap volumes in P1, for example, decreased mean dose differences with P4 from 20.2-13.3Gy, respectively, while correcting individual swallowing muscles for the higher PTV_E doses in P2 reduced mean dose differences by 4.2-6.7Gy. For the further investigation, only corrected values are considered as these provide a fairer comparison between the periods.

Table 3 shows significant associations between mean dose and period for all OARs (except IL-SMG), demonstrating that advances made throughout the periods were associated with improved OAR sparing. Compared to IMRT (P1), RapidArc (P2) was associated with averaged mean dose reductions to the CL / IL-parotids of 2.4Gy / 5.2Gy. For the CL-SMG, introduced in P2, significant improvements in sparing (9.9Gy) were also obtained. Neither swallowing muscle nor oral cavity sparing was routinely performed in P1-P2, resulting in similar doses.

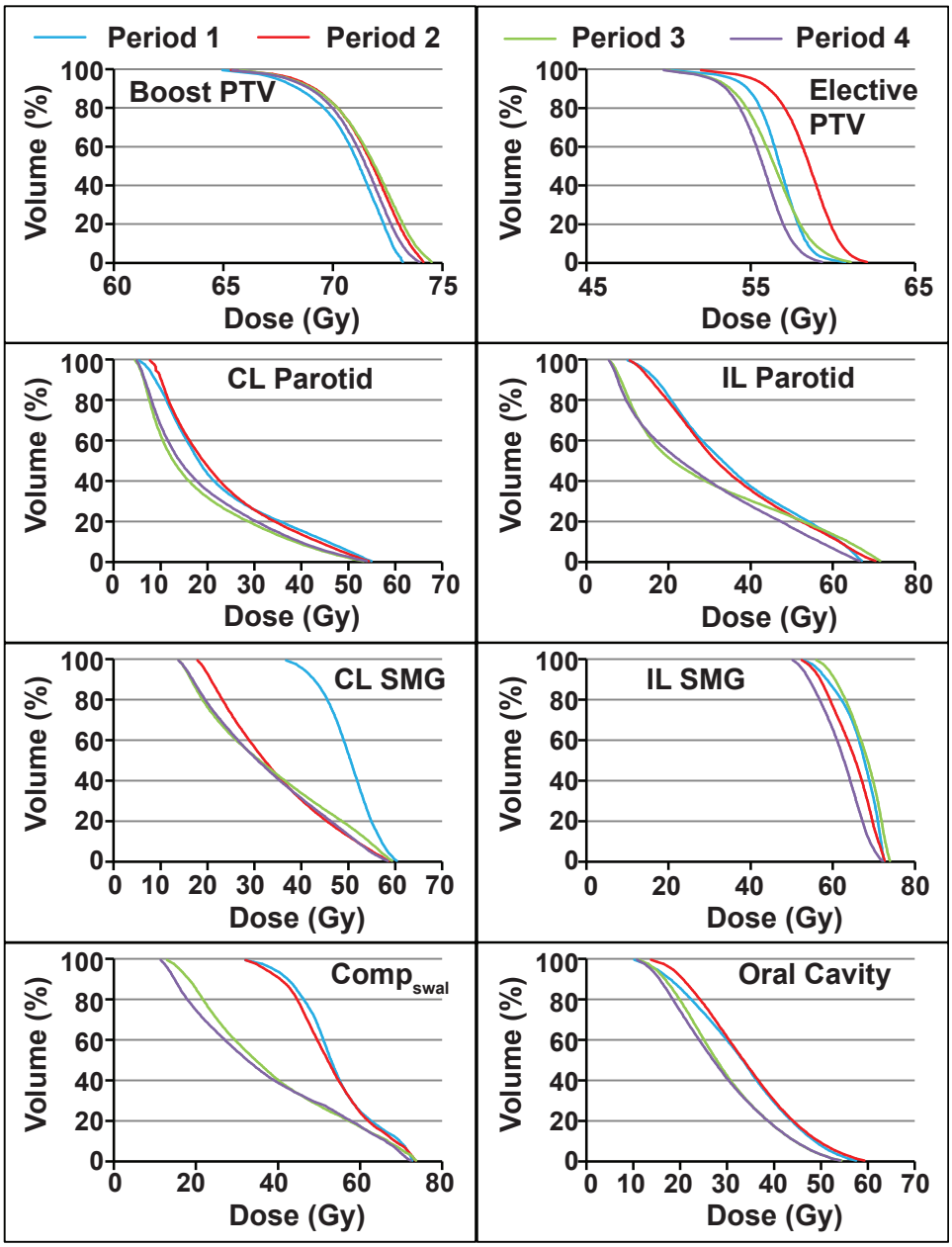


Figure 2. Averaged dose-volume histogram (DVH)-lines for a number of structures showing the progressive improvements in the subsequent periods.

Table 3. Averaged mean dose differences and 95% confidence interval (in brackets) of all organs-at-risk (OARs) with respect to period 4. The values that were corrected for differences between the periods are shown in bold, with the footnotes stating which factors were corrected for. Positive values indicate improvements in period 4.

Period	1	2	3	p-values			
	Diff. (95% CI)	Diff. (95% CI)	Diff. (95% CI)	overall[a]	P1 vs P2[b]	P2 vs P3[b]	P3 vs P4[b]
Salivary Glands							
CL parotid[c]	7.9 (3.3 to 12.5) **5.0 (2.3 to 7.7)**	4.8 (0.2 to 9.4) **2.6 (0 to 5.3)**	-2.6 (-7.2 to 2.1) **-0.9 (-3.6 to 1.7)**	$<10^{-3}$ **$<10^{-3}$**	0.43 **0.099**	$<10^{-3}$ **0.003**	0.83 **1.00**
IL parotid	13.8 (6.5 to 21.1)	8.6 (1.3 to 15.9)	3.0 (-4.3 to 10.3)	$<10^{-3}$	0.35	0.25	1.00
CL SMG[c]	20.2 (13.1 to 27.2) **13.3 (7.3 to 19.3)**	3.3 (-4.1 to 10.7) **3.4 (-2.4 to 9.2)**	2.1 (-4.8 to 8.9) **1.1 (-4.3 to 6.5)**	$<10^{-3}$ **$<10^{-3}$**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**	1.00 **1.00**
IL SMG	2.7 (-1.1 to 6.6)	2.4 (-1.3 to 6.1)	3.5 (-0.5 to 7.4)	0.092	1.00	1.00	0.12
$Comp_{sal}$	9.9 (5.3 to 14.5)	6.5 (1.8 to 11.1)	-0.4 (-4.9 to 4.2)	$<10^{-3}$	0.30	0.001	1.00
Swallowing Muscles							
Cricopharyngeal Muscle[d]	24.0 (18.3 to 29.6) **22.4 (16.8 to 28.1)**	26.8 (21.1 to 32.4) **22.6 (15.9 to 29.3)**	2.4 (-3.3 to 8.0) **2.4 (-3.1 to 7.9)**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**
Lower Larynx[d]	27.8 (21.5 tot 34.2) **25.3 (19.3 to 31.4)**	28.5 (22.2 to 34.9) **21.8 (14.6 to 29.0)**	5.2 (-1.1 to 11.6) **5.2 (-0.7 to 11.1)**	$<10^{-3}$ **$<10^{-3}$**	1.00 **0.87**	$<10^{-3}$ **$<10^{-3}$**	0.17 **0.11**
Upper Larynx[d]	13.1 (4.1 to 22.2) **11.4 (2.2 to 20.6)**	14.0 (5.0 to 23.1) **9.4 (-1.5 to 20.2)**	3.3 (-5.8 to 12.3) **3.3 (-5.6 to 12.1)**	$<10^{-3}$ **0.008**	1.00 **1.00**	0.01 **0.80**	1.00 **1.00**
PCM[g] Inferior[d]	16.3 (8.7 to 24.0) **14.3 (6.6 to 22.0)**	17.8 (10.1 to 25.5) **12.2 (2.9 to 21.5)**	-3.9 (-11.5 to 3.8) **-3.9 (-11.3 to 3.5)**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**	$<10^{-3}$ **$<10^{-3}$**	1.00 **0.98**
PCM[g] Medial[f]	3.7 (-2.8 to 10.2) **3.9 (-0.5 to 8.2)**	4.6 (-1.9 to 11.1) **5.4 (1.1 to 9.6)**	-1.0 (-7.4 to 5.5) **4.1 (-0.3 to 8.6)**	0.061 **0.007**	1.00 **1.00**	0.14 **1.00**	1.00 **0.082**
PCM[g] Superior[c]	5.0 (0-0.8 to 10.8) **3.2 (-0.4 to 6.8)**	5.0 (-0.9 to 10.8) **4.0 (0.4 to 7.6)**	-0.7 (-6.5 to 5.1) **3.3 (-0.4 to 6.9)**	0.009 **$<10^{-3}$**	1.00 **1.00**	0.060 **1.00**	1.00 **0.11**
UES[h]	24.6 (18.8 to 30.5)	28.3 (22.4 to 34.2)	4.4 (-1.5 to 10.2)	$<10^{-3}$	0.60	$<10^{-3}$	0.28
$Comp_{swal}$[d]	18.1 (12.6 to 23.6) **16.5 (11.0 to 21.9)**	19.8 (14.3 to 25.3) **15.3 (8.7 to 21.9)**	3.8 (-1.6 to 9.3) **3.8 (-1.4 to 9.1)**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**	$<10^{-3}$ **$<10^{-3}$**	0.38 **0.32**
Oral Cavity[e]	7.6 (2.1 to 13.2) **8.9 (3.7 to 14.1)**	10.3 (4.7 to 15.9) **8.9 (3.6 to 14.2)**	1.5 (-4.1 to 7.0) **1.7 (-3.5 to 6.8)**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**	$<10^{-3}$ **$<10^{-3}$**	1.00 **1.00**

[a] Overall p-value for mean dose differences between the periods

[b] p-value between specific periods, correct for multiple testing (Bonferroni, times 6)

[c] Corrected for differences in overlap between organ and planning target volume (PTV)

[d] Corrected for higher elective PTV doses in periods 1 and 2

[e] Corrected for differences in organ volume

[f] Corrected for both overlap volume and organ volume

[g] Pharyngeal constrictor muscle

[h] Upper esophageal sphincter

Swallowing muscle sparing (P3) resulted in significant mean dose reductions to 4 out of 7 swallowing muscles and $comp_{swal}$, with average differences of 11.5-23.9Gy with respect to P2. Despite these improvements, salivary gland sparing also improved in P3, with significantly improved CL-parotid / $comp_{sal}$ mean doses (7.4Gy / 6.9Gy). Routine oral cavity sparing also commenced in P3, resulting in significant mean dose reductions (7.2Gy) compared to P2. P3 and P4 plans were similar with no significant differences between the OARs, although $comp_{swal}$ and oral cavity sparing improved by 3.8Gy and 1.7Gy, respectively, in P4. A significant decrease in required monitor units was found between P1 (1050 on average) and P2 / P3 / P4 (432 / 492 / 516). For the VMAT plans (P2-P4), this shows that increased modulation was associated with improved quality.

Table 4 shows PTV dose coverage and homogeneity values and various dose deposition metrics. When necessary, the latter were corrected for higher PTV_E prescription doses in P1-P2 and differences in PTV size between the periods. Table 4 shows that improved OAR sparing in P3-P4 did not decrease PTV dose homogeneity or increase dose deposition in the body. Conversely, the best PTV dose homogeneity values were obtained in P4, with improvements of 0.1%-0.7% / 0%-2.5% over HI_B / HI_E in P1-P3. Boost CI slightly improved in P3-P4 compared to P2, although elective CI got slightly worse in P4. Comparable mean dose values to the boost and elective ring structures were obtained in P2-P4, while Body – PTV doses decreased for a number of parameters in P3-P4 compared to P2. The absolute differences between P2-P4 are small, and unlikely to be clinically relevant. However, the IMRT plans (P1) had larger differences several in dose deposition metrics compared to the RapidArc plans, which may have contributed to the statistical significance.

To evaluate the effect of different dose calculation algorithms throughout the periods, AAA v10.0.28 was used to recalculate three plans from P1 / P4, originally calculated using PBCA v8.1.14 / Acuros v11.0.31. Compared to PBCA, AAA improved HI_B / HI_E at most by 1.0% / 1.1%, while D_{sal} / D_{swal} /D_{oc} remained within 0.4Gy / 0.9Gy / 0.3Gy of their original values. Acuros and AAA plans were similar, with HI_B / HI_E values remaining within 0.3% / 0.4%, and D_{sal} / D_{swal} / D_{oc} within 0.4Gy / 0.6Gy / 0.4Gy.

Table 4. Average values and 95% confidence intervals (in brackets) of the planning target volume (PTV) dose coverage and homogeneity values and dose deposition metrics in all periods. The values that were corrected for differences between the periods, with respect to period 4, are shown in bold, with the footnotes stating which factors were corrected for.

Period	1	2	3	4	p-value[a]
Boost PTV					
V95[d] (%)	98.1 (97.5 to 98.7)	98.8 (98.4 to 99.1)	99.0 (98.8 to 99.2)	98.9 (98.9 to 99.0)	0.047
V107[d] (%)	0.4 (0 to 0.8)	1.3 (0.6 to 2.0)	2.8 (1.2 to 4.5)	0.6 (0.3 to 0.9)	0.001
HI_B[e] (%)	9.8 (9.5 to 10.2)	10.0 (9.5 to 10.5)	10.4 (10.0 to 10.8)	9.7 (9.3 to 9.9)	0.064
Elective PTV					
V95 (%)	98.5 (98.1 to 98.9)	98.2 (97.8 to 98.5)	97.7 (97.3 to 98.1)	98.2 (98.0 to 98.5)	0.017
V107 (%)	7.7 (5.7 to 9.7)	12.3 (8.2 to 16.5)	15.3 (12.3 to 18.4)	7.6 (5.8 to 9.5)	$< 10^{-3}$
HI_E[e] (%)	13.3 (12.5 to 14.1)	13.6 (12.8 to 14.4)	15.8 (15.0 to 16.6)	13.3 (12.7 to 13.9)	$< 10^{-3}$
Conformity Indices					
Boost (%)	1.10 (1.07 to 1.12)	1.17 (1.12 to 1.22)	1.14 (1.12 to 1.16)	1.14 (1.11 to 1.16)	0.018
Elective (%)	1.56 (1.51 to 1.61)	1.39 (1.37 to 1.42)	1.35 (1.32 to 1.37)	1.42 (1.38 to 1.45)	$< 10^{-3}$
Mean dose to boost PTV ring structures					
0cm (Gy)	54.4 (52.6 to 56.3) **54.0 (51.3 to 56.9)**	53.8 (52.0 to 55.7) **52.4 (50.4 to 54.5)**	52.7 (50.9 to 54.6) **52.4 (50.8 to 54.2)**	52.7 (52.0 to 53.5)	0.004 **$< 10^{-3}$ [b]**
1cm (Gy)	39.1 (35.7 to 42.5) **38.9 (35.6 to 42.1)**	38.8 (35.4 to 42.2) **38.3 (35.1 to 41.6)**	39.8 (36.3 to 43.1) **39.2 (36.0 to 42.5)**	38.6 (37.2 to 40.0)	0.715 **$< 10^{-3}$ [c]**
2cm (Gy)	32.8 (29.2 to 36.4) **35.2 (29.5 to 40.8)**	32.0 (28.5 to 35.6) **30.4 (26.4 to 34.5)**	32.0 (28.4 to 35.6) **31.6 (28.2 to 35.0)**	31.7 (30.2 to 33.2)	0.739 **$< 10^{-3}$ [b]**
Mean dose to elective PTV ring structures					
0cm (Gy)	47.7 (46.3 to 49.1) **46.9 (45.0 to 48.8)**	46.9 (45.6 to 48.3) **44.9 (43.6 to 46.3)**	44.7 (43.3 to 46.1) **44.5 (43.3 to 45.6)**	44.7 (44.2 to 45.2)	$< 10^{-3}$ **$< 10^{-3}$ [b]**
1cm (Gy)	31.1 (29.7 to 32.4) **31.5 (29.5 to 33.4)**	30.1 (28.8 to 31.5) **28.9 (27.5 to 30.3)**	29.0 (27.6 to 30.3) **28.7 (27.6 to 29.9)**	29.0 (28.5 to 29.5)	$< 10^{-3}$ **$< 10^{-3}$ [b]**
2cm (Gy)	23.0 (21.6 to 24.4) **25.1 (23.0 to 27.2)**	23.6 (22.2 to 25.0) **22.9 (21.4 to 24.4)**	22.8 (21.4 to 24.2) **22.7 (21.4 to 23.9)**	22.9 (22.4 to 23.4)	0.334 **$< 10^{-3}$ [b]**
Body-PTV volume					
Mean Dose (Gy)	20.5 (19.0 to 22.0) **21.1 (18.9 to 23.3)**	20.6 (19.1 to 22.1) **19.6 (18.0 to 21.2)**	18.7 (17.2 to 20.2) **18.5 (17.2 to 19.9)**	18.9 (18.3 to 19.5)	$< 10^{-3}$ **$< 10^{-3}$ [b]**
V5Gy[f] (%)	73.8 (69.7 to 59.9) **77.4 (70.8 to 66.0)**	79.1 (75.0 to 65.2) **78.7 (74.0 to 65.5)**	74.5 (70.4 to 60.6) **74.0 (70.1 to 60.0)**	73.5 (71.7 to 57.3)	$< 10^{-3}$ **$< 10^{-3}$ [b]**
V30Gy[f] (%)	29.7 (26.3 to 33.2) **30.7 (25.4 to 36.1)**	29.6 (26.2 to 33.1) **27.2 (23.4 to 31.0)**	27.2 (23.8 to 30.7) **26.7 (23.6 to 29.9)**	27.3 (26.0 to 28.7)	0.018 **$< 10^{-3}$ [b]**
V50Gy[f] (%)	7.4 (5.9 to 9.0) **7.5 (5.5 to 9.5)**	7.0 (5.5 to 8.6) **4.9 (3.5 to 6.4)**	4.5 (3.0 to 6.1) **4.2 (3.1 to 6.3)**	5.3 (4.8 to 5.9)	$< 10^{-3}$ **$< 10^{-3}$ [b]**

[a] Overall p-value for mean dose differences between the periods

[b] Corrected for higher elective PTV doses in periods 1 and 2 and differences in combined PTV size

[c] Corrected for differences in combined PTV size

[d] V95% / V107%: PTV Volumes receiving 95% / 107% of the prescribed dose

[e] HI_B / HI_E: Boost / elective PTV homogeneity indices

[f] V5Gy / V30Gy / V50Gy: Body – PTV volumes receiving 5Gy / 30Gy / 50Gy

Discussion

The present study provided a longitudinal evaluation of changes in HNC plans resulting from the introduction of new radiotherapy technologies, planning techniques and increasing experience. Advances made throughout the periods significantly associated with mean OAR dose reductions. Statistical corrections were applied to account for differences in geometric and dosimetric composition of the periods. We are unaware of studies that performed a similar analysis on routine clinical HNC plans. Closest was an investigation by Chen et al. whether evolving treatment modalities over thirty years associated with improved clinical outcomes for base-of-tongue carcinoma patients[27].

Our analysis confirmed that improvements reported in previous studies[19,21,22] have been translated into routine clinical care. Although the number of actively spared OARs increased considerably over time, this did not increase dose to initially spared OARs nor increase dose deposition in the remainder of the body. OAR sparing improved substantially over the years, with average improvements of 9.9Gy / 16.5Gy / 8.9Gy for D_{sal} / D_{swal} / D_{oc}, when comparing the first (manually optimized IMRT plans, 2005-2008) and last periods (automatically optimized VMAT plans, 2014-2015). Gains in OAR sparing did not come at the cost of PTV dose coverage and homogeneity values or increased dose deposition in the body – PTV. The plans from earlier periods were likely less optimal, allowing such improvements without obvious detriment elsewhere. To illustrate this, the first ten patients from P2 were manually replanned following the interactive optimization approach described in our current planning protocol for HNC. As shown by Table 5 , this resulted in substantial improvements in OAR sparing compared to P2, with relatively similar PTV dose homogeneity values, conformity indices and body – PTV doses. Although only oropharynx cancer patients were included in the present study, similar gains are anticipated for other sites since the advances in dose delivery and treatment planning techniques have been implemented for the entire head and neck region.

In line with previous research[21], IMRT (P1) and initial dual-arc VMAT plans (P2) were comparable in terms of PTV doses, OAR sparing and dose deposition. Large mean dose decreases were noted for both parotids in P3, suggesting that increased planner skill and experience in RapidArc planning played an important role in improved OAR sparing, aided by more effective planning protocols and additional training. A number of recent studies comparing IMRT with VMAT illustrate that, if planner skill and approach to planning are similar, reasonably comparable plan quality can be obtained using both techniques[28–30]. If the IMRT plans from P1 would therefore be replanned using current planning approaches, similar plan quality as in P4 will likely be obtained.

After geometric correction, minor improvements in D_{swal} / D_{oc} were obtained in P4 with respect to P3 (3.8Gy / 1.7Gy). This was expected based on previous work[19], showing that AIO could improve OAR sparing without decreasing PTV dose homogeneity or dose conformity, or increasing dose deposition. Reductions in OAR doses can be interpreted using radiobiological models to estimate normal tissue complication probabilities (NTCPs). For example, using Figure 3 in the paper by Dijkema et al.[31], if complication is defined as a stimulated parotid gland flow rate to <25% of pre-treatment, the NTCP could at maximum be lowered by 20% by reducing parotid gland mean doses by approximately 7Gy (comparable to P1 versus P3-P4). Further analysis is merited to determine whether the longitudinal dosimetric improvements translate into better patient reported outcome measures[32], which was beyond the scope of the present study.

Statistical correction was necessary to account for geometric and dosimetric differences between the periods. Since we evaluated the dosimetric results of clinically delivered plans, this is inherent to this type of study. We attempted to create homogeneous groups by only selecting patients with the primary tumor originating in the tonsillar region and pharyngeal wall. A number of potential limitations should be noted; (i) Only 30 patients were selected for each period, while the inclusion of more patients could provide a more robust statistical analysis. (ii) Although no correction was necessary for differences in combined PTV size between the periods for the OARs, contributions of separate PTVs were not evaluated, even though Table 2 showed significant differences between some periods. (iii) Overcorrection of OAR doses was noted; Correcting the swallowing muscles for PTV_E doses resulted in 4.2-6.7Gy mean dose decreases, while actual differences in prescribed doses were only 3.5Gy (57.75Gy – 54.25Gy).

In conclusion, the introduction of new radiotherapy technologies and planning techniques for locally advanced head and neck cancer patients at our department over the last decade was associated with substantially improved OAR sparing while providing similar PTV dose coverage and homogeneity. These gains did not degrade conformity indices or increase dose deposition in the body. This is important to consider when comparing photon plans with other new irradiation techniques (e.g. protons).

Table 5. Results of the first 10 patients of period 2 (P2), replanned using the treatment planning protocols of P3-P4. The data shows that the new approaches to treatment planning improved all considered metrics, without obvious detriment elsewhere. This was despite the fact that the boost PTV volumes were significantly larger in P2 compared to P3-P4. Significant differences between the plans, determined using two-sided Student t-tests with significance level set at $p \leq 0.05$, are indicated by the '*' symbol.

Plan	Clinical P2	Replanned P2	# Included
Boost planning target volume			
V95% (%)	99.1 ± 1.0	98.8 ± 0.8	10
V107% (%)	0.5 ± 0.9	0.5 ± 0.7	10
HI_B (%)	9.0 ± 1.1	9.6 ± 1.0	10
Elective planning target volume			
V95% (%)	98.8 ± 0.9	98.9 ± 0.5	10
V107% (%)	9.5 ± 8.6	4.0 ± 3.6	10
IH_E (%)	12.5 ± 2.0	11.4 ± 1.6	10
Conformity Indices			
Boost PTV (%)	1.13 ± 0.10	1.18 ± 0.11	10
Elective PTV (%)	1.41 ± 0.07	1.36 ± 0.19	10
Max Dose (Gy)			
Spinal Cord	44.5 ± 6.7	41.3 ± 4.1*	10
Brainstem	43.4 ± 6.8	38.9 ± 3.7*	7
Mean Dose (Gy)			
CL Parotid	24.1 ± 5.7	18.7 ± 5.8*	10
IL Parotid	33.6 ± 10.5	28.7 ± 10.7*	10
CL SMG	32.4 ± 3.5	27.3 ± 2.4*	5
Trachea	41.2 ± 0.2	38.1 ± 4.0	2
Esophagus	41.0 ± 1.7	34.9 ± 6.6	2
Body-PTV volume			
Mean dose (Gy)	20.4 ± 2.4	20.7 ± 2.0	10
V5Gy (%)	78.3 ± 3.9	80.7 ± 3.0*	10
V30Gy (%)	28.6 ± 6.0	28.3 ± 4.7	10
V50Gy (%)	7.0 ± 2.8	7.5 ± 2.2	10
Monitor Units	430 ± 42	481 ± 50*	10

References

1 A Eisbruch, L. A Dawson, H.M. Kim, C.R. Bradford, J.E. Terrell, D.B. Chepeha, T.N. Teknos, Y. Anzai, L.H. Marsh, M.K. Martel, R.K. ten Haken, G.T. Wolf, and J.A. Ship, "Conformal and intensity modulated irradiation of head and neck cancer: the potential for improved target irradiation, salivary gland function, and quality of life.," Acta Otorhinolaryngol. Belg. **53**(3), 271–5 (1999).

2 H.P. van der Laan, A. Gawryszuk, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, E.W. Korevaar, O. Chouvalova, K. Wopken, H.P. Bijl, and J.A. Langendijk, "Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction.," Radiother. Oncol. **107**(3), 282–7 (2013).

3 C.M. Nutting, J.P. Morden, K.J. Harrington, T.G. Urbano, S.A. Bhide, C. Clark, E.A. Miles, A.B. Miah, K. Newbold, M. Tanay, F. Adab, S.J. Jefferies, C. Scrase, B.K. Yap, R.P. A'Hern, M.A. Sydenham, M. Emson, and E. Hall, "Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial.," Lancet Oncol. **12**(2), 127–36 (2011).

4 E.H.N. Pow, D.L.W. Kwong, A.S. McMillan, M.C.M. Wong, J.S.T. Sham, L.H.T. Leung, and W.K. Leung, "Xerostomia and quality of life after intensity-modulated radiotherapy vs. conventional radiotherapy for early-stage nasopharyngeal carcinoma: initial report on a randomized controlled clinical trial.," Int. J. Radiat. Oncol. Biol. Phys. **66**(4), 981–91 (2006).

5 G. Sanguineti, M.P. Sormani, S. Marur, G.B. Gunn, N. Rao, M. Cianchetti, F. Ricchetti, T. McNutt, B. Wu, and A.A. Forastiere, "Effect of radiotherapy and chemotherapy on the risk of mucositis during intensity-modulated radiation therapy for oropharyngeal cancer.," Int. J. Radiat. Oncol. Biol. Phys. **83**(1), 235–42 (2012).

6 A. Eisbruch, H.M. Kim, F.Y. Feng, T.H. Lyden, M.J. Haxer, M. Feng, F.P. Worden, C.R. Bradford, M.E. Prince, J.S. Moyer, G.T. Wolf, D.B. Chepeha, and R.K. ten Haken, "Chemo-IMRT of oropharyngeal cancer aiming to reduce dysphagia: swallowing organs late complication probabilities and dosimetric correlates.," Int. J. Radiat. Oncol. Biol. Phys. **81**(3), e93–9 (2011).

7 K. Wopken, H.P. Bijl, A. van der Schaaf, H.P. van der Laan, O. Chouvalova, R.J.H.M. Steenbakkers, P. Doornaert, B.J. Slotman, S.F. Oosting, M.E.M.C. Christianen, B.F.A.M. van der Laan, J.L.N. Roodenburg, C.R. Leemans, I.M. Verdonck-de Leeuw, and J.A. Langendijk, "Development of a multivariable normal tissue complication probability (NTCP) model for tube feeding dependence after curative radiotherapy/chemo-radiotherapy in head and neck cancer.," Radiother. Oncol. **113**(1), 95–101 (2014).

8 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F. a M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

9 American Society for Radiation Oncology, ASTRO Model Policies - Intensity Modulated Radiation Therapy (IMRT) (2015).

10 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

11 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

12 S. Everitt, T. Kron, N. Fimmell, J. Reynolds, C. Laferlita, D. Ball, M. Schneider-Kolsky, R. Budd, and M. Macmanus, "Interplanner variability in carrying out three-dimensional conformal radiation therapy for non-small-cell lung cancer.," J. Med. Imaging Radiat. Oncol. **52**(3), 293–6 (2008).

13 Y. Matsuo, K. Takayama, Y. Nagata, E. Kunieda, K. Tateoka, N. Ishizuka, T. Mizowaki, Y. Norihisa, M. Sakamoto, Y. Narita, S. Ishikura, and M. Hiraoka, "Interinstitutional variations in planning for stereotactic body radiation therapy for lung cancer.," Int. J. Radiat. Oncol. Biol. Phys. **68**(2), 416–25 (2007).

14 R.G.J. Kierkels, R. Visser, H.P. Bijl, J.A. Langendijk, A.A. van 't Veld, R.J.H.M. Steenbakkers, and E.W. Korevaar, "Multicriteria optimization enables less experienced planners to efficiently produce high quality treatment plans in head and neck cancer radiotherapy.," Radiat. Oncol. **10**(1), 87 (2015).

15 J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, "Evaluation of a knowledge-based planning solution for head and neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **91**(3), 612–20 (2015).

16 E.M. Quan, J.Y. Chang, Z. Liao, T. Xia, Z. Yuan, H. Liu, X. Li, C.A. Wages, R. Mohan, and X. Zhang, "Automated volumetric modulated Arc therapy treatment planning for stage III lung cancer: how does it compare with intensity-modulated radio therapy?," Int. J. Radiat. Oncol. Biol. Phys. **84**(1), e69–76 (2012).

17 S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

18 P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, A. Al-Mamgani, L. Incrocci, and B.J.M. Heijmen, "Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients.," Int. J. Radiat. Oncol. Biol. Phys. **88**(5), 1175–9 (2014).

19 J.P. Tol, M. Dahele, J. Peltola, J. Nord, B.J. Slotman, and W.F.A.R. Verbakel, "Automatic interactive optimization for volumetric modulated arc therapy planning," Radiat. Oncol. **10**(1), 75 (2015).

20 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–35 (2011).

21 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–9 (2009).

22 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, "Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.," Radiat. Oncol. **6**(1), 74 (2011).

23 M.E.M.C. Christianen, J.A. Langendijk, H.E. Westerlaan, T.A. van de Water, and H.P. Bijl, "Delineation of organs at risk involved in swallowing for radiotherapy treatment planning.," Radiother. Oncol. **101**(3), 394–402 (2011).

24 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, "Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity," Med. Phys. **41**(2), 021722 (2014).

25 H.C. de Boer, and B.J. Heijmen, "A protocol for the reduction of systematic patient setup errors with minimal portal imaging workload, " Int. J. Radiat. Oncol. Biol. Phys. **50**(5),1350–65 (2001).

26 C. Thilmann, A. Zabel, S. Nill, B. Rhein, A. Hoess, P. Haering, S. Milke-Zabel, W. Harms, W. Schlegel, M. Wannenmacher, and J. Debus, "Intensity-modulated radiotherapy of the female breast.," Med. Dosim. **27**(2), 79–90 (2002).

27 L.A. Chen, C.J. Anker, J.P. Hunt, L.O. Buchmann, K.F. Grossmann, K. Boucher, L.M.C. Fang, D.C. Shrieve, and Y.J. Hitchcock, "Clinical outcomes associated with evolving treatment modalities and radiation techniques for base-of-tongue carcinoma: thirty years of institutional experience.," Cancer Med. **4**(5), 651–60 (2015).

28 D. Van Gestel, C. van Vliet-Vroegindeweij, F. Van den Heuvel, W. Crijns, A. Coelmont, B. De Ost, A. Holt, E. Lamers, Y. Geussens, S. Nuyts, D. Van den Weyngaert, T. Van den Wyngaert, J.B. Vermorken, and V. Gregoire, "RapidArc, SmartArc and TomoHD compared with classical step and shoot and sliding window intensity modulated radiotherapy in an oropharyngeal cancer treatment plan comparison.," Radiat. Oncol. **8**(1), 37 (2013).

29 W. Lechner, G. Kragl, and D. Georg, "Evaluation of treatment plan quality of IMRT and VMAT with and without flattening filter using Pareto optimal fronts.," Radiother. Oncol. **109**(3), 437–41 (2013).

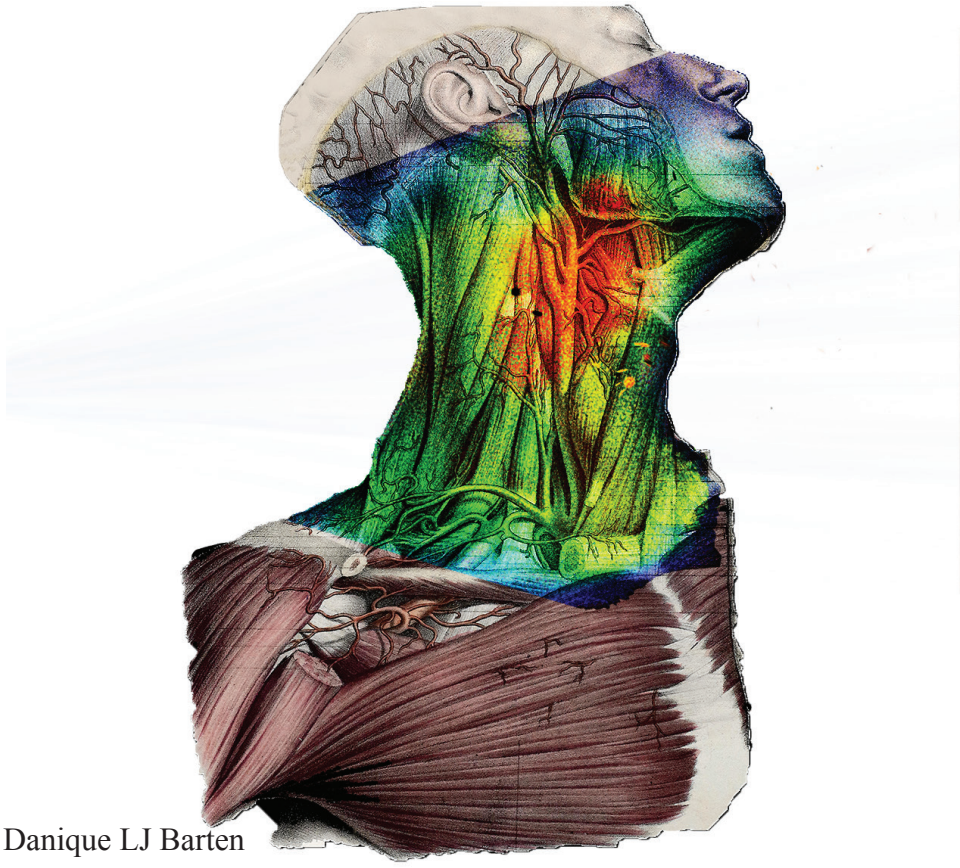
30 T. Wiezorek, T. Brachwitz, D. Georg, E. Blank, I. Fotina, G. Habl, M. Kretschmer, G. Lutters, H. Salz, K. Schubert, D. Wagner, and T.G. Wendt, "Rotational IMRT techniques compared to fixed gantry IMRT and tomotherapy: multi-institutional planning study for head-and-neck cases.," Radiat. Oncol. **6**(1), 20 (2011).

31 T. Dijkema, C.P.J. Raaijmakers, R.K. ten Haken, J.M. Roesink, P.M. Braam, A.C. Houweling, M.A. Moerland, A. Eisbruch, and C.H.J. Terhaard, "Parotid gland function after radiotherapy: the combined michigan and utrecht experience.," Int. J. Radiat. Oncol. Biol. Phys. **78**(2), 449–53 (2010).

32 R.N. Rinkel, I.M. Verdonck-de Leeuw, P. Doornaert, J. Buter, R. de Bree, J.A. Langendijk, N.K. Aaronson, and C.R. Leemans, "Prevalence of swallowing and speech problems in daily life after chemoradiation for head and neck cancer based on cut-off scores of the patient-reported outcome measures SWAL-QOL and SHI.," Eur. Arch. Otorhinolaryngol. **273**(7), 1849-55 (2015).

Chapter 11

Comparison of organ-at-risk sparing and plan robustness for spot-scanning proton therapy and volumetric modulated arc photon therapy in head-and-neck cancer



Danique LJ Barten
Jim P Tol
Max Dahele
Ben J Slotman
Wilko FAR Verbakel

Medical Physics **42**(11), 6589-6598 (2015).

Abstract

Purpose

Proton radiotherapy for head and neck cancer (HNC) aims to improve organ-at-risk (OAR) sparing over photon radiotherapy. However, it may be less robust for setup and range uncertainties. We investigated OAR sparing and plan robustness for spot-scanning proton planning techniques and compared these with volumetric modulated arc therapy (VMAT) photon plans.

Materials and Methods

10 HNC patients were replanned using dual-arc VMAT (RapidArc) and spot-scanning proton techniques. Included OARs were the contralateral and ipsilateral parotid and submandibular glands and individual swallowing muscles. Proton plans were made using Multi-Field Optimization (MFO, using 3, 5 and 7 fields) and Single-Field Optimization (SFO, using 3 fields). OAR sparing was evaluated using mean dose to composite salivary glands ($comp_{sal}$) and composite swallowing muscles ($comp_{swal}$). Plan robustness was determined for setup errors and range uncertainties (±3mm and ±3% HU, respectively) evaluating V95% and V107% for clinical target volumes (CTVs).

Results

Averaged over all patients $comp_{sal}$ / $comp_{swal}$ mean doses were lower for the 3-field MFO plans (14.6Gy / 16.4Gy) compared to the 3-field SFO plans (20.0Gy / 23.7Gy) and VMAT plans (23.0Gy / 25.3Gy). Using more than 3 fields resulted in differences in OAR sparing of less than 1.5Gy between plans. SFO plans were significantly more robust than MFO plans, while the VMAT plans were the most robust.

Conclusions

MFO plans had improved OAR sparing but were less robust than SFO and VMAT plans, while SFO plans were more robust than MFO plans but resulted in less OAR sparing. Robustness of the MFO plans did not increase with more fields.

Introduction

Radiation therapy is effective and commonly used for the treatment of head and neck cancer (HNC). However, it may result in persistent and troublesome side effects, including xerostomia and dysphagia[1–3]. Dose reductions to the salivary glands and swallowing muscles may lower the risk and / or severity of these complications[1–4]. For example, Deasy et al.[1] have shown that the parotid gland mean dose is linearly associated with late toxicity. New radiation techniques such as intensity modulated radiotherapy (IMRT), or its rotational variant volumetric modulated arc therapy (VMAT), have improved organ-at-risk (OAR) sparing in the past decade[5,6]. Proton radiotherapy has the potential to further improve OAR sparing because the finite range of the protons causes a steep dose fall-off after its dose maximum (the 'Bragg Peak').

In spot-scanning proton therapy (SSPT), narrow proton beams are moved by scanning magnetic fields over the transverse plane to position the Bragg Peak in the target volume[7]. The beam energy is modified to control the dose deposition depth of the proton beam. If the minimum beam energy is too high to cover shallow target regions, the depth of the Bragg Peak can be further reduced by inserting a set of range shifter plates. This decreases the effective beam energy and consequently reduces the Bragg Peak depth[8]. However, a range shifter also broadens the beam in the air gap between the nozzle and the patient resulting in an increased spot size[7–9]. An alternative way of reducing the beam energy is to use a bolus or build-up region closely applied to the patient's surface[10].

Proton therapy dose distributions can be shaped using either Single-Field optimization (SFO) or Multi-Field optimization (MFO). In SFO, also known as single-Field uniform dose (SFUD), individual optimization of each field results in a uniform dose distribution that provides adequate target coverage per field[11]. In MFO, typically referred to as intensity modulated proton therapy (IMPT), the spots of each field are optimized simultaneously and only the combined fields deliver adequate and homogeneous target dose coverage[11]. Since MFO allows for the creation of more complex dose distributions than SFO (analogous to comparing IMRT to 3D conformal radiotherapy), it can further improve OAR sparing.

Several previous studies have investigated OAR sparing using MFO compared to different photon techniques[12–14] and SFO[15] for the treatment of complex HNC patients. These studies demonstrate that, given the complexity of HNC patients, MFO can improve OAR sparing over both photon radiotherapy and SFO. Other research has shown that reducing the proton beam spot size may further improve OAR sparing[16,17]. However, with the exception of Quan et al.[15], none of these studies evaluated OAR sparing together with plan robustness (i.e. the invariance of the dose distribution to uncertainties in planning and delivery).

In photon radiotherapy, a planning target volume (PTV) margin is used to ensure adequate coverage of the CTV when positioning errors occur during the treatment course. Because the physical dose distributions of photons are relatively invariant under setup uncertainties[18], possible regions of under dosage are commonly located near the borders of the PTV. However, this may not hold true for MFO proton plans since the steep dose fall-off after the Bragg Peak in combination with inhomogeneous individual fields make these plans highly sensitive to uncertainties in anatomical changes, patient positioning (setup errors) and variations in the density / Hounsfield Units (HU) of the planning CT-scan (range uncertainties)[19–22]. SFO plans are expected to be more robust to uncertainties than MFO plans, because they utilize less modulation of the Bragg Peak intensities of each field[23]. Since setup and range uncertainties may cause significant deviations between the planned and delivered dose distributions it is important to evaluate robustness of proton therapy plans. In a robust plan, deviations in setup and range uncertainties have little effect on dose delivered to the target volumes. Neglecting the analysis of plan robustness when comparing different proton plans, or when comparing proton plans with photon plans, may result in unfair comparisons, since both the achievable OAR sparing and target dose coverage may be affected by the plan robustness. Several studies have discussed uncertainties in MFO treatments[19–21,24,25] and designed robust optimization methods[22,26–31]. These studies often aim to validate in-house developed robust optimization algorithms, which are not available in commercial proton planning systems. For this reason, we have evaluated the robustness of MFO and SFO SSPT plans made with a clinically available treatment planning system without a robust optimization algorithm, and quantified the loss of dose coverage of the CTV under conditions of different setup and range uncertainties.

The aforementioned studies highlight the interest in dosimetric comparisons between different SSPT techniques and photon VMAT / IMRT, and the robustness of SSPT techniques. However, they have typically not simultaneously investigated OAR sparing and robustness or provided a comprehensive investigation of the influence of different proton delivery techniques (e.g. MFO or SFO) and the number of fields on OAR sparing and robustness. In addition, VMAT planning techniques have also improved[32,33]. The focus of the present study was therefore to perform a systematic and detailed investigation of the plan quality and robustness of a range of different SSPT techniques and compare these against the latest clinically implemented VMAT techniques. Locally advanced head and neck cancer plans that included sparing of the salivary glands, oral cavity and individual swallowing muscles were used for this analysis.

Materials and Methods

Patient group

This study included 10 arbitrarily selected patients who had previously received RapidArc (Varian Medical Systems, Palo Alto, USA) VMAT treatment for HNC at our department between 2012 and 2013. Disease site, stage and planning target volumes (PTV) are shown in Table 1.

Target volumes and organs-at-risk

All plans were created using our institutional simultaneous integrated boost (SIB) technique[5,34] delivering 54.25Gy to the elective PTV (PTV_E) and 70Gy to the boost PTV (PTV_B) in 35 fractions of 1.55Gy and 2Gy, respectively. PTV_B consisted of a 4-5mm expansion of the CTV, which was defined as a 5mm expansion of the gross tumor volume and biopsy proven positive lymph node regions. PTV_E consisted of a similar expansion of the elective nodal regions, minus a 5mm transition PTV (PTV_T) that was created between PTV_B and PTV_E to facilitate a steep dose fall off between them. PTV expansions were 4mm above the shoulders and 5mm below them. The same PTVs were used for both the proton and photon plans. The plans aimed to deliver 95% of the prescribed dose to 99% / 98% of PTV_B / PTV_E, respectively (V95% > 99% / 98%), while PTV_B / PTV_E volumes receiving more than 107% should be smaller than 4% / 15% (V107% < 4% / 15%).

Table 1. Disease site, stage, and volumes of boost, transition and elective planning target volumes (PTV_B, PTV_T and PTV_E, respectively) and the volumes of the composite salivary glands ($comp_{sal}$) and composite swallowing ($comp_{swal}$) muscles for all patients.

Patient number	Disease Site	Stage	PTV_B (cm^3)	PTV_T (cm^3)	PTV_E (cm^3)	$comp_{sal}$ (cm^3)	$comp_{swal}$ (cm^3)
1	Oropharynx	T2N0	70.2	28.3	346.6	41.7	17.3
2	Oropharynx	T3N2c	313.0	103.0	433.8	74.0	17.3
3	Oropharynx	T2N2a	94.5	43.4	240.6	42.4	25.0
4	Oropharynx	T2N2b	188.6	48.3	441.7	82.3	16.5
5	Oropharynx	T2N2b	289.0	34.4	457.9	70.4	18.5
6	Oropharynx	T4aN1	164.5	79.5	280.8	78.9	32.0
7	Oropharynx	T4aN1	143.7	62.9	360.4	73.0	35.0
8	Supraglottic larynx	T3N2b	143.4	80.1	358.5	42.4	9.9
9	Supraglottic larynx	T2N1	61.9	30.1	247.5	68.5	23.2
10	Hypopharynx	T3N2b	182.7	64.1	416.4	47.4	7.4

Salivary structures included in the optimization could consist of the ipsilateral and contralateral parotid and submandibular glands. The swallowing muscles could consist of the upper esophageal sphincter, upper and lower larynx, superior, medial and inferior pharyngeal constrictor muscle, cricopharyngeal muscle and the esophagus. The treating radiation oncologist could decide not to spare OARs that had a large overlap with the PTVs. For OAR dose reporting, composite salivary glands ($comp_{sal}$) and swallowing muscles ($comp_{swal}$) consisting of combined, individually optimized OAR, were created. The size of these structures is shown in Table 1.

The first priority for OARs during optimization was to keep the point doses for the spinal cord and brainstem well below their respective dose tolerance levels. Under these constraints, the mean dose to the salivary glands and swallowing muscles was reduced as much as possible while satisfying our institutional criteria for PTV dose coverage and homogeneity, described previously. To prevent dose hotspots outside the target regions, maximum dose objectives were applied on normal tissues and ring structures that were created around both PTVs. If hotspots (dose > 65Gy) appeared outside the PTVs, these volumes were delineated and assigned lower maximum dose constraints in a subsequent plan optimization. All plans were normalized to provide the same mean PTV_B dose as the original clinical photon plan, which typically fell between 101-103% of the prescribed dose.

Treatment planning techniques

The clinical VMAT plans, using two full arcs and 6MV photons, were created in the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, USA) and optimized using the progressive resolution optimizer (PRO, v10.0.28). Dose calculation was performed using the anisotropic analytical algorithm (AAA, v10.0.28) with a 2.5mm calculation grid. A 'continue previous optimization' was performed after the AAA calculation to improve PTV dose homogeneity[32]. Optimization was performed following our previously described institutional planning protocol[5,6,32,34]. In brief, this uses 3 to 5 optimization objectives for the individual salivary and swallowing structures. During optimization, the aim was to try and maintain a fixed diagonal distance between the various OAR optimization objectives and their respective dose-volume histogram (DVH)-lines that are displayed in the Eclipse optimization window. Using this method, progressive dose reductions to the OARs are attempted throughout the interactive optimization process, while the planning constraints on the PTVs ensure that adequate target coverage is maintained.

Proton treatment plans were created in Eclipse using the proton convolution superposition (PCS, v11.0.30) algorithm with a 2.5mm dose calculation grid. Spot spacing was automatically selected by the PCS algorithm depending on the spot size at each beam depth. For example, the spot spacing was selected as 6.3mm for a 4mm spot size in air. Proton beam data was provided by Varian Medical Systems as a standard SSPT data set, measured at a clinical facility using Varian equipment.

For each patient 3, 5 and 7-field MFO plans and 3-field SFO plans were created. Gantry angles were chosen based on the general shape of the PTVs and previous studies[13]. The 3-field SFO and MFO plans used gantry angles of 30°, 180° and 330°. In the 5-field MFO plans, gantry angles of 150° and 210° were added. Gantry angles of 30°, 100° / 110°, 150°, 180°, 210° / 230°, 260° and 330° were used in the 7-field plans depending on the position of the parotid glands and shoulders. Target margins were chosen 0.1cm proximal, 0.3cm distal and 0.3cm lateral to the target volume for each beam direction. PTV-based optimization was used to create the proton plans because a robust optimization algorithm was not available in our proton planning system. Plan optimization was performed similar to the interactive optimization used to create the VMAT plans. Per patient, 1 to 3 optimizations were performed for each plan. The best plan, judged by the level of PTV dose coverage and the absence of hotspots, was selected for further analysis.

The proton beam energy ranged from 70MeV to 210MeV. Since a 70MeV proton beam penetrates approximately 4.0cm in water, a range shifter was needed to cover proximal parts of the target volume. A range shifter of 5.7cm water equivalent material was used in this study. However, a range shifter introduces scatter and broadens the proton beam spot size[10] due to the air gap between the range shifter plates and the patient. As an alternative to the range shifter, we therefore also evaluated the use of a build-up region (bolus) placed directly outside the patient's body in order to achieve adequate dose coverage of the proximal target volume and a relatively smaller spot size. Initial evaluations showed that sufficient target coverage could be achieved using a 3cm build-up region. Build-up (BU) and range shifter (RS) plans were created for both MFO and SFO techniques to further analyze the influence of spot size on OAR sparing and plan robustness.

An air gap of 2cm between the snout and the nearest patient surface was used for all plans. The air gap between snout and patient at the central axis therefore differs per field, plan and patient.

Plan Evaluation

VMAT and proton plans were evaluated on the basis of OAR sparing (mean dose to individual and composite OARs), PTV dose coverage and robustness to interfraction setup errors (SE) and range uncertainties (RU). SE were simulated for all plans by shifting the isocenter -3mm and +3mm (all possible combinations) in three directions (anterior-posterior, superior-inferior and right-left directions). Finally, the worst-case scenarios[35,36], in the form of simultaneous shifts of 3mm in all directions, are reported. The influence of RU on the proton plans was investigated by introducing -3% and +3% errors in planning-CT HU values (consistent with Moyers et al.[37]), resulting in over- and undershoot of the Bragg Peak location. SE and RU were also combined (SRU, 4 combinations) to evaluate the plan robustness of worst-case scenarios. Intrafraction errors were not considered in this study.

SE were simulated for both proton and photon plans. The negligible effect of RU on photon plan quality was demonstrated for two VMAT plans, and SRU was not considered separately for the VMAT plans. Plan robustness to SE, RU and SRU was evaluated on the basis of boost and elective clinical target volumes (CTVs) receiving at least 95% (V95%) and 107% (V107%) of the prescribed dose. Paired two-sided Student t-tests were performed to investigate whether the differences in plan robustness were statistically significant ($p \leq 0.05$) for SE, RU and SRU between i) MFO, SFO and VMAT plans, and ii) MFO plans using a different number of fields.

In this planning study, the physical dose is considered and this has not been corrected for radiobiological effects.

To summarize, the following comparisons were considered in the present study:

1. PTV dose coverage for MFO (3 / 5 / 7-fields), SFO (3-fields) and VMAT.
2. OAR mean doses for MFO (3 / 5 / 7-fields), SFO (3-fields) and VMAT.
3. Plan robustness (±3mm setup errors (SE), ±3% HU variations (RU) and 4 combinations of SE and RU) for MFO (3 / 5 / 7-fields), SFO (3-fields) and VMAT (only ±3mm setup errors).

Results

Build-up versus range shifter

For all patients, the use of build-up region instead of a range shifter reduced mean dose to $comp_{sal}$ and $comp_{swal}$. Averaged over all MFO plans, these reductions were 2.6Gy and 4.4Gy, respectively. A reduction in mean OAR dose was not unexpected because of the larger

proton beam spot size when using a range shifter due to (i) the increased ranger shifter thickness (5.7cm compared to 3.0cm for the build-up region), and (ii) the distance that the protons have to travel in the air gap between the range shifter plates and patient. When investigating the difference in spot sizes between build-up and range shifter, however, we found an unexpectedly large effect on the spot size for beams with low energies (≤110MeV). The spot sigma in air at isocenter of the range shifter proton beam was 9.9mm and 4.4mm for the 100MeV and 210MeV beams, respectively (with default setting of range shifter at 2cm distance from the body surface) compared to 4.5mm and 4.0mm when using a build-up region. After discussing this with Varian Medical Systems, it was discovered that when the beam data is measured with the range shifter at a large distance from the phantom surface, as was the case with the beam data that had been provided, the beam configuration requires extrapolations over large distances. This was performed inaccurately for low proton energies. The excessive extrapolations in the provided beam data only affected the range shifter beams with energies ≤110MeV. This is likely to have affected the results of range shifter beams attempting to cover the proximal part of the patient, leading to higher OAR doses. These findings have led Varian to design a new model that correctly handles the scatter effects due to the range shifter, regardless of its position. As this corrected prototype was not yet available, we have chosen not to include further range shifter data in this paper and only the results of the build-up plans will be further presented and discussed.

Target volume coverage

Table 2 summarizes the dosimetric results, averaged over all patients. Clinically acceptable maximum doses to the spinal cord and brainstem were achieved in all plans. PTV_B / PTV_E dose coverage and homogeneity was acceptable for all MFO and VMAT plans, whereas this was not always true for the SFO plans. For 2 SFO BU plans adequate dose coverage of PTV_B was not obtained (mean V95% of 97.2±0.1%). The aim of maintaining PTV_E V107% < 15% was not satisfied for the SFO BU plans of 3 patients (mean over all patients of 14.6%±4.2%). To illustrate the wide range of OAR sparing and PTV dose coverage and homogeneity values that were obtained, Figure 1 shows the DVH-lines of all plans that were created for patient 1.

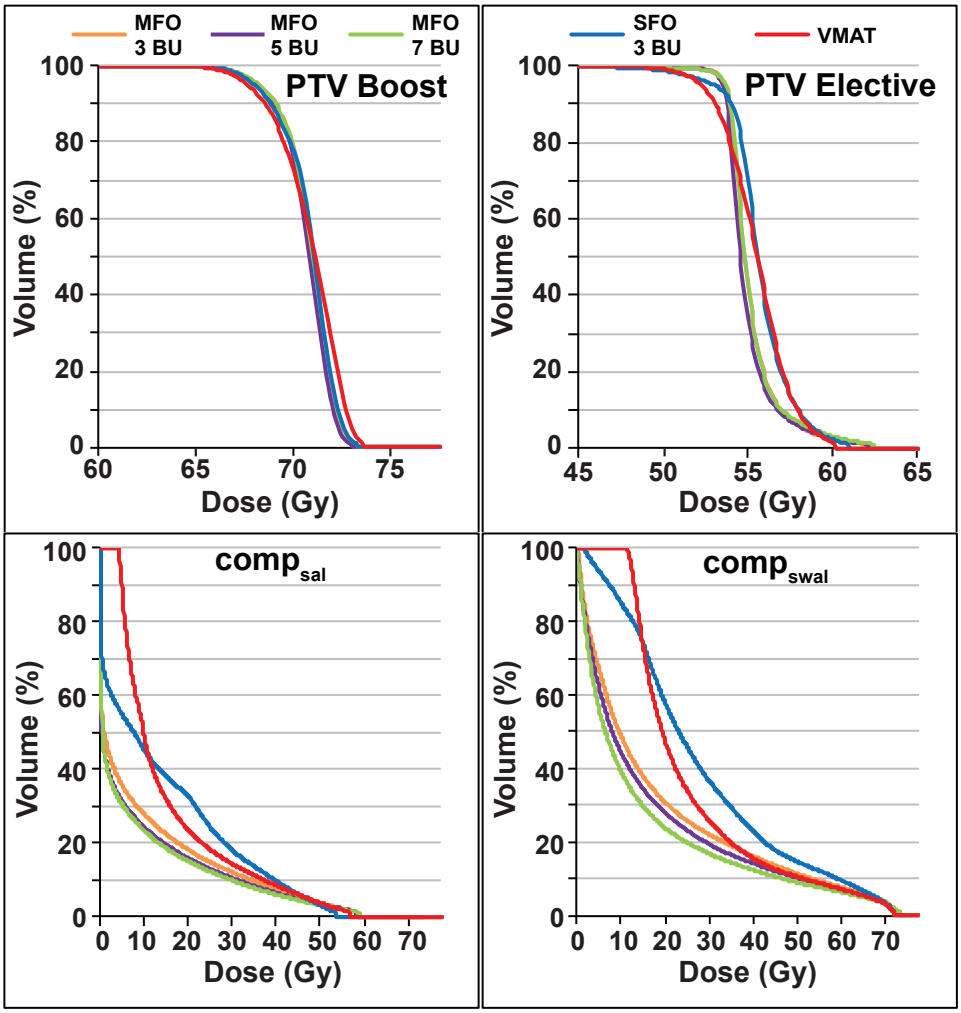


Figure 1. Dose-volume histograms (DVHs) for all plans of patient 1. Upper figures: DHVs for the boost planning target volume (PTV_B, left) and elective PTV (PTV_E, Right). Lower figures: DVHs for the composite salivary glands ($comp_{sal,}$ left) and composite swallowing muscles ($comp_{swal}$, right). MFO 3 / 5 / 7 BU refers to Multi-field Optimization using 3 / 5 / 7 fields and a 3cm build-up region.

Organ-at-risk sparing

Comparing MFO with SFO and VMAT

Averaged over all patients, the lowest mean doses to the salivary glands and swallowing muscles were achieved in the MFO plans (see Table 2).

Number of fields

Averaged over all patients, increasing the number of fields only marginally decreased the mean dose to the salivary glands and swallowing muscles (1.5Gy / 1.2Gy for $comp_{sal}$ / $comp_{swal}$ at most).

Plan robustness

Comparing MFO with SFO and VMAT

As an example, Figure 2 shows the impact of all individual uncertainties on deteriorating the CTV_B and CTV_E coverage for all plans of patient 1. One of the worst case scenario's, the combination of +3mm SE and +3% RU, resulted in a CTV_B V95% of only 79.9% for the MFO 7-field BU plan, while this was 87.9% for the 3-field SFO BU plan, the most robust proton BU plan.

SFO plans were found to be significantly more robust than MFO plans. Averaged over all patients and 4 SRU, the 3-field SFO BU plans resulted in V95% of CTV_B / CTV_E > 98.3%. VMAT plans were the most robust with V95% of CTV_B / CTV_E values of 99.7% / 99.3% averaged over all patients, although only SE were considered. Since this did not include the RU, we calculated the impact of a simulation of CT-scan HU changes of +40HU and -40HU for 2 patients. This showed variations in PTV_B and CTV_B V95% no greater than 0.5% and 0%, respectively.

Figure 3 shows the averaged effect of SE, RU and SRU on CTV_B V95% for the plans of all 10 patients. In contrast to Figure 2, where V95% of CTVs is shown for all errors individually, in Figure 3 the 2 SE, 2 RU and 4 SRU are averaged per plan for each patient. This shows that SRU results in the lowest CTV_B dose coverage values on average.

Figure 4 shows the impact of the individual uncertainties on CTV_B and CTV_E V95% and V107% of the 3-field MFO BU plans, for all patients. For example, CTV_B V95% was <98% in 6 patients for SRU +3mm and +3%HU while this was the case for 8 patients for SRU -3mm and +3%HU.

Number of fields

Increasing the number of MFO fields did generally not result in higher CTV dose coverage values. Figure 5 shows the impact of the combined uncertainty (-3mm and +3% HU) on CTV_B / CTV_E V95% for all plans of all patients.

In general, MFO plans with additional fields did not improve CTV_B and CTV_E V95% under conditions of setup / range uncertainties (Figure 2, 3 and 5). The CTV_E V107% in the build-up plans decreased by at most 0.8%, with increasing number of fields.

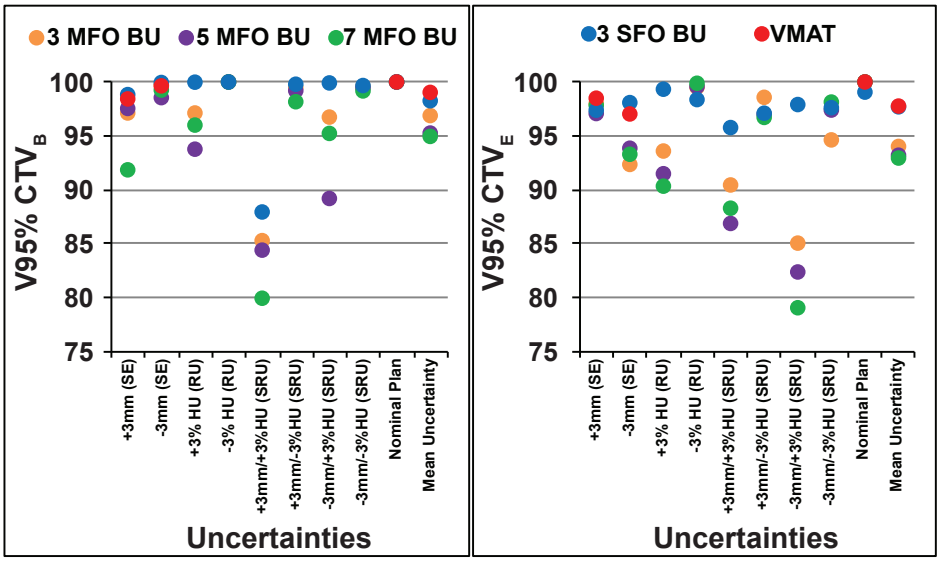


Figure 2. Impact of the uncertainties on V95% of the boost clinical target volume (CTV_B, left) and elective CTV (CTV_E, right) for patient 1 are shown for all proton plans and the VMAT plan. MFO 3 / 5 / 7 BU refers to Multi-Field Optimization using 3 / 5 / 7 fields and a 3cm build up region.

Table 2. Dosimetric results and mean doses of relevant structures for VMAT and proton plans, averaged over all 10 patients. Ranges are given between the parentheses.

	PHOTONS	PROTONS			
	VMAT[a]	MFO[b] Build-up			SFO[c] Build-up
	2 Arcs	3 Fields	5 Fields	7 Fields	3 Fields
Boost planning target volume (%)					
V95%	99.2 (98.6 - 99.8)	99.3 (98.7 - 100.0)	99.4 (98.8 - 99.9)	99.4 (98.3 - 100.0)	98.8 (97.2 - 99.9)
V107%	1.9 (0.0 - 8.8)	0.6 (0.0 - 2.9)	0.5 (0.0 - 1.8)	0.4 (0.0 - 1.7)	0.7 (0.0 - 2.3)
Elective planning target volume (%)					
V95%	97.7 (97.1 - 99.1)	99.4 (99.0 - 99.7)	99.5 (99.2 - 99.8)	99.5 (99.2 - 99.8)	98.6 (96.4 - 99.6)
V107%	13.3 (5.3 - 22.5)	7.3 (4.7 - 9.4)	5.5 (3.1 - 7.6)	5.5 (0.0 - 4.7)	14.6 (9.1 - 23.4)
Mean dose (Gy)					
IL Parotid[d]	25.0 (14.3 - 33.7)	17.5 (8.1 - 25.5)	16.9 (6.9 - 24.3)	16.7 (7.0 - 23.8)	21.3 (11.0 - 28.3)
CL Parotid[d]	19.2 (11.0 - 29.6)	10.3 (4.0 - 21.7)	9.7 (3.4 - 22.6)	9.2 (3.2 - 21.2)	14.0 (5.8 - 26.0)
CL SMG[e]	28.9 (23.0 - 35.9)	21.9 (11.9 - 33.7)	20.5 (11.3 - 32.9)	20.1 (10.4 - 32.5)	34.9 (23.3 - 41.3)
$comp_{sal}$[f]	23.0 (15.2 - 31.7)	14.6 (9.2 - 22.0)	13.8 (8.1 - 21.0)	13.5 (7.8 - 21.5)	20.0 (14.0 - 26.5)
$comp_{swal}$[g]	25.3 (17.0 - 40.5)	16.8 (4.5 - 36.0)	16.5 (4.6 - 35.9)	16.4 (5.1 - 35.9)	23.7 (10.1 - 36.9)
Max dose (Gy)					
Spinal Cord	40.8 (38.8 - 41.8)	32.9 (28.6 - 36.4)	33.7 (32.3 - 39.1)	33.3 (32.1 - 36.4)	31.7 (24.8 - 38.5)

[a] Volumetric modulated arc therapy

[b] Multi-Field optimization

[c] Single-Field optimization

[d] Ipsilateral and contralateral parotid glands

[e] Contralateral submandibular gland

[f] Composite salivary glands

[g] Composite swallowing muscles

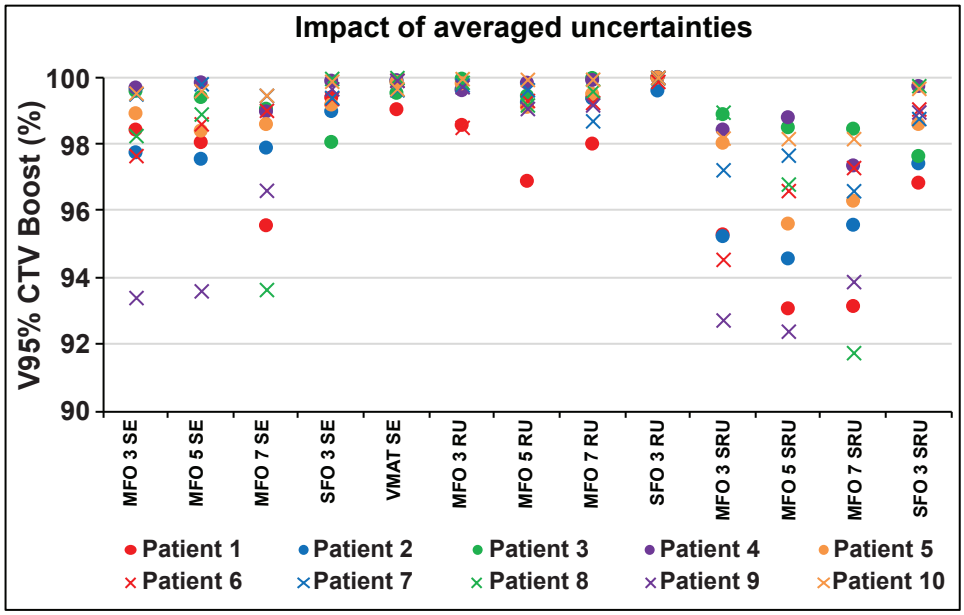


Figure 3. For all patients the impact of 2 averaged setup errors (SE, left), 2 averaged range uncertainties (RU, middle) and 4 averaged combinations of both (SRU, right) on the V95% of the Boost CTV per plan. MFO 3 / 5 / 7 refers to Multi-Field Optimization using 3 / 5 / 7 fields. SFO 3 refers to Single-Field Optimization using 3 fields.

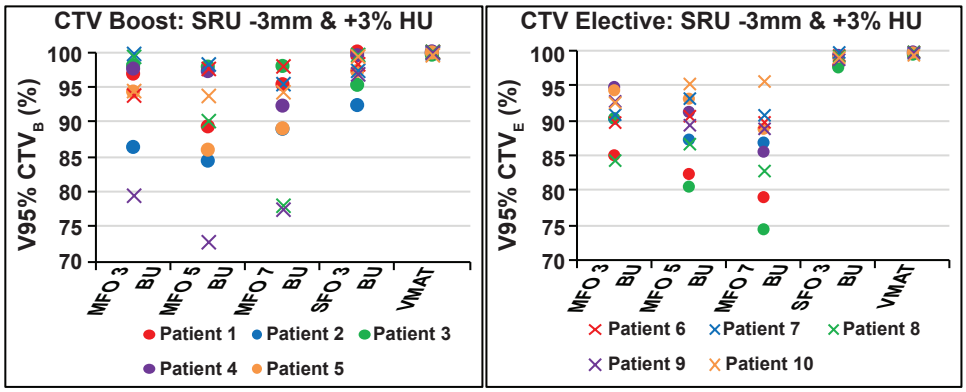


Figure 5. For all patients the impact of -3mm set-up errors and +3% HU range uncertainties on V95% of the Boost CTV (left) and V95% of the Elective CTV (right). In the VMAT plans only -3mm set-up errors are taken into account. MFO 3 / 5 / 7 refers to Multi-Field Optimization using 3 / 5 / 7 fields. SFO 3 refers to Single-Field Optimization using 3 fields.

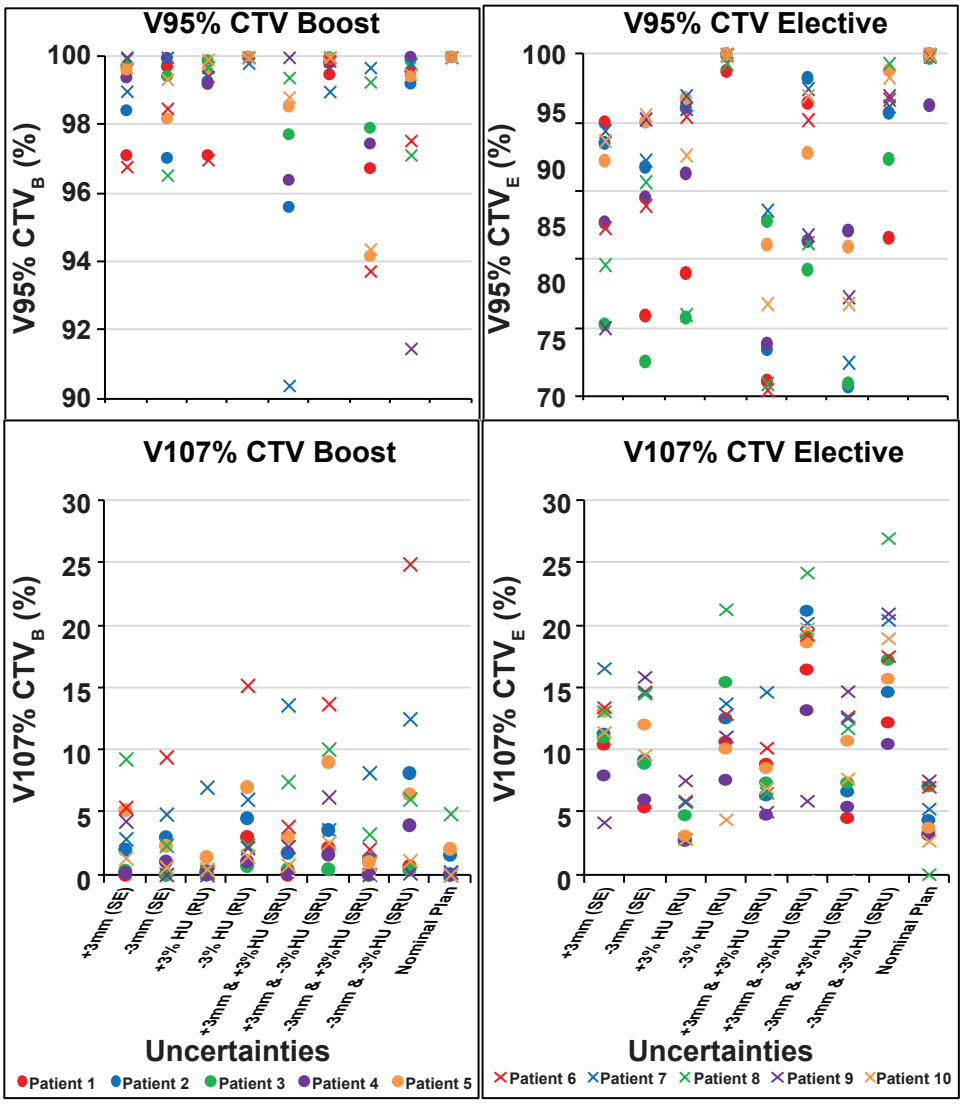


Figure 4. For all patients, the impact of set-up errors and range uncertainties on the Boost and Elective CTVs for the 3-field MFO proton plan.

Discussion

This study presented results of a treatment planning comparison of both SFO and MFO spot-scanning proton delivery techniques and VMAT photon radiotherapy for HNC, and evaluated OAR sparing and plan robustness for plans generated on a commercial treatment planning system without a robust optimization algorithm.

Organ-at-risk sparing

Comparing MFO with SFO and VMAT

Consistent with previous research[12–15], our results showed that MFO improved sparing of the salivary glands and swallowing muscles over SFO and VMAT. However, since SFO attempts to deliver a homogeneous target dose with every field, the chosen field orientations might play an important role in achievable OAR sparing. In this study, the same field orientations were used for the 3-field SFO and MFO plans. A higher contralateral submandibular gland mean dose was generally obtained in the SFO plans because this structure was in the beam's eye view with beam angles of 30° or 330°. We suspect that changing beam angle orientations may influence the resulting plan more than, for example, modifying the optimization strategy with the same beam orientations. Optimal field arrangements are likely patient specific and different between SFO and MFO techniques, however, investigating this was beyond the scope of the present study.

Number of fields

The present results show that increasing the number of fields improved $comp_{sal}$ and $comp_{swal}$ mean doses by 1.2Gy and 1.5Gy, respectively, averaged over all patients. Van der Laan et al.[13] found that increasing the number of MFO fields from 3 to 7 did not improve salivary gland sparing while the mean dose to individual swallowing muscles generally improved by 1-2Gy (except for a 6Gy mean dose reduction for the supraglottic larynx). Similarly, Lomax et al.[7] and Steneker et al.[16] reported little gain in OAR sparing by increasing the number of fields, although no quantitative data was provided. Furthermore, only a single target region was used in the study of Steneker et al.[16] and no attempt was made to spare the swallowing muscles. In contrast, our study dealt with multiple PTV structures and focused on both salivary gland (primary) and swallowing muscle (secondary) sparing, increasing the complexity of treatment planning. Although analyzing the normal tissue complication probabilities (NTCPs) was beyond the scope of the present study, previous reports that investigated the potential gain in OAR sparing using MFO[12–14] predicted clinical benefits for dysphagia and xerostomia.

Plan robustness

Comparing MFO with SFO and VMAT

VMAT plans were found to be the most robust and, in line with previous research[15,23], SFO plans were more robust than MFO plans for all uncertainties (SE, RU and SRU). Evaluating individual proton plans showed large variation in robustness over the patients. CTV_B V95%, for example, ranged 79.0-99.0% for the 3-field MFO BU plans. This may in part be attributed to differences in tissue heterogeneity along the beam direction of the patients.

Number of fields

Although Unkelbach et al.[21] suggested that additional fields may improve robustness of proton plans our findings are similar to those of Kraan et al.[25], who found that additional fields did not improve MFO plan robustness. Chosen field orientations play an important role in plan robustness because proton beams passing through heterogeneous tissue tend to be more sensitive to uncertainties than beams passing through homogeneous media due to degradation of the sharp dose fall-off of the Bragg Peak[19]. However, the large target volumes and complex OAR-PTV geometries in the head and neck region make it difficult to select beam orientations that provide a proper balance between plan quality, OAR sparing and plan robustness when using a non-robust optimization algorithm.

Limitations

Robust optimization

A potential limitation of this study is that robust optimization was not available in the proton planning system, and the use of conventional PTV margins may not be optimal for MFO[11,28,31,38]. In future investigations it would be desirable to use robust optimization algorithms or probabilistic planning solutions that simultaneously strive for plan robustness and OAR sparing during optimization[21,26,29,30]. The available literature regarding the influence of robust optimization algorithms on OAR sparing reported only marginally increased OAR doses for robustly optimized plans[27,28].

In addition, we have not quantitatively evaluated mean OAR doses resulting from SE, RU and SRU, as the treatment planning system was not able to report this data, nor did it allow export of the perturbed OAR DVHs. In analysis of plan robustness, effects of intrafraction shifts (i.e., patient movement between delivering the different beams) and anatomical changes / deformations with respect to the planning CT were not investigated. With current online imaging techniques, the risk of ±3mm setup errors in 3 directions simultaneously is

small, especially in the target region where proper patient positioning is most important. However, taking into account the residual setup inaccuracies, deformations, possible intrafraction shifts and anatomical changes that occur over the course of treatment, the investigated setup errors no longer present an unrealistic scenario for evaluating plan robustness.

Further research is needed to evaluate influence of patient specific beam orientations, uncertainties due to intrafraction shifts and anatomical changes, robust optimization algorithms and differences between planning systems. In light of patient specific differences in OAR sparing and plan robustness, identifying which specific patients / groups of patients are likely to benefit most from proton therapy is necessary[39].

This study demonstrated that intensity modulated proton therapy without robust optimization can lead to undesirable CTV underdosing when setup and / or range uncertainties are present. MFO plans provided the lowest mean doses to the salivary glands and swallowing muscles but were less robust compared to SFO and VMAT. Although SFO improved plan robustness, OAR sparing was only marginally better than VMAT. Increasing the number of MFO fields marginally improved OAR sparing and did not improve plan robustness. Future studies should compare VMAT with robustly optimized proton plans to reveal the additional value of MFO for head and neck radiotherapy.

References

1 J.O. Deasy, V. Moiseenko, L. Marks, K.S.C. Chao, J. Nam, and A. Eisbruch, "Radiotherapy dose-volume effects on salivary gland function.," Int. J. Radiat. Oncol. Biol. Phys. **76**(3 Suppl), S58–63 (2010).

2 H.P. van der Laan, A. Gawryszuk, M.E.M.C. Christianen, R.J.H.M. Steenbakkers, E.W. Korevaar, O. Chouvalova, K. Wopken, H.P. Bijl, and J.A. Langendijk, "Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: treatment planning optimization and clinical introduction.," Radiother. Oncol. **107**(3), 282–7 (2013).

3 M. Little, M. Schipper, F.Y. Feng, K. Vineberg, C. Cornwall, C.A. Murdoch-Kinch, and A. Eisbruch, "Reducing Xerostomia After Chemo-IMRT for Head-and-Neck Cancer: Beyond Sparing the Parotid Glands," Int. J. Radiat. Oncol. **83**(3), 1007–14 (2011).

4 M.E.M.C. Christianen, C. Schilstra, I. Beetz, C.T. Muijs, O. Chouvalova, F.R. Burlage, P. Doornaert, P.W. Koken, C.R. Leemans, R.N.P.M. Rinkel, M.J. de Bruijn, G.H. de Bock, J.L.N. Roodenburg, B.F.A.M. van der Laan, B.J. Slotman, I.M. Verdonck-de Leeuw, H.P. Bijl, and J.A. Langendijk, "Predictive modelling for swallowing dysfunction after primary (chemo)radiation: results of a prospective observational study.," Radiother. Oncol. **105**(1), 107–14 (2012).

5 P. Doornaert, W.F.A.R. Verbakel, M. Bieker, B.J. Slotman, and S. Senan, "RapidArc planning and delivery in patients with locally advanced head-and-neck cancer undergoing chemoradiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **79**(2), 429–435 (2011).

6 W.F.A.R. Verbakel, J.P. Cuijpers, D. Hoffmans, M. Bieker, B.J. Slotman, and S. Senan, "Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study.," Int. J. Radiat. Oncol. Biol. Phys. **74**(1), 252–259 (2009).

7 A.J. Lomax, T. Böhringer, A. Bolsi, D. Coray, F. Emert, G. Goitein, M. Jermann, S. Lin, E. Pedroni, H. Rutz, O. Stadelmann, B. Timmermann, J. Verwey, and D.C. Weber, "Treatment planning and verification of proton therapy using spot scanning: initial experiences.," Med. Phys. **31**(11), 3150–3157 (2004).

8 E. Pedroni, S. Scheib, T. Böhringer, A. Coray, M. Grossmann, S. Lin, and A. Lomax, "Experimental characterization and physical modelling of the dose distribution of scanned proton pencil beams.," Phys. Med. Biol. **50**(3), 541–561 (2005).

9 U. Titt, D. Mirkovic, G.O. Sawakuchi, L.A. Perles, W.D. Newhauser, P.J. Taddei, and R. Mohan, "Adjustment of the lateral and longitudinal size of scanned proton beam spots using a pre-absorber to optimize penumbrae and delivery efficiency.," Phys. Med. Biol. **55**(23), 7097–7106 (2010).

10 S. Safai, T. Bortfeld, and M. Engelsman, "Comparison between the lateral penumbra of a collimated double-scattered beam and uncollimated scanning beam in proton radiotherapy.," Phys. Med. Biol. **53**(6), 1729–50 (2008).

11 S.E. McGowan, N.G. Burnet, and A.J. Lomax, "Treatment planning optimisation in proton therapy.," Br. J. Radiol. **86**(1021), 20120288 (2013).

12 S. Kandula, X. Zhu, A.S. Garden, M. Gillin, D.I. Rosenthal, K.K. Ang, R. Mohan, M. V Amin, J.A. Garcia, R. Wu, N. Sahoo, and S.J. Frank, "Spot-scanning beam proton therapy vs intensity-modulated radiation therapy for ipsilateral head and neck malignancies: A treatment planning comparison.," Med. Dosim. **38**(4), 1–5 (2013).

13 H.P. van der Laan, T.A. van de Water, H.E. van Herpt, M.E.M.C. Christianen, H.P. Bijl, E.W. Korevaar, C.R. Rasch, A.A. van 't Veld, A. van der Schaaf, C. Schilstra, and J.A. Langendijk, "The potential of intensity-modulated proton radiotherapy to reduce swallowing dysfunction in the treatment of head and neck cancer: A planning comparative study," Acta Oncol. (Madr). **52**(3), 1–9 (2012).

14 T.A. van de Water, A.J. Lomax, H.P. Bijl, M.E. de Jong, C. Schilstra, E.B. Hug, and J.A. Langendijk, "Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer.," Int. J. Radiat. Oncol. Biol. Phys. **79**(4), 1216–1224 (2011).

15 E.M. Quan, W. Liu, R. Wu, Y. Li, S.J. Frank, X. Zhang, X.R. Zhu, and R. Mohan, "Preliminary evaluation of multifield and single-field optimization for the treatment planning of spot-scanning proton therapy of head and neck cancer.," Med. Phys. **40**(8), 081709 (2013).

16 M. Steneker, A. Lomax, and U. Schneider, “Intensity modulated photon and proton therapy for the treatment of head and neck tumors.,” Radiother. Oncol. **80**(2), 263–267 (2006).

17 T.A. van de Water, A.J. Lomax, H.P. Bijl, C. Schilstra, E.B. Hug, and J.A. Langendijk, “Using a reduced spot size for intensity-modulated proton therapy potentially improves salivary gland-sparing in oropharyngeal cancer.,” Int. J. Radiat. Oncol. Biol. Phys. **82**(2), e313–9 (2012).

18 D. Maleike, J. Unkelbach, and U. Oelfke, “Simulation and visualization of dose uncertainties due to interfractional organ motion.,” Phys. Med. Biol. **51**(9), 2237–52 (2006).

19 A.J. Lomax, “Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties.,” Phys. Med. Biol. **53**(4), 1027–1042 (2008).

20 A.J. Lomax, “Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions.,” Phys. Med. Biol. **53**(4), 1043–1056 (2008).

21 J. Unkelbach, T. Bortfeld, B.C. Martin, and M. Soukup, “Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning.,” Med. Phys. **36**(1), 149–163 (2009).

22 D. Pflugfelder, J.J. Wilkens, and U. Oelfke, “Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy.,” Phys. Med. Biol. **53**(6), 1689–1700 (2008).

23 F. Albertini, E.B. Hug, and A.J. Lomax, “Is it necessary to plan with safety margins for actively scanned proton therapy?,” Phys. Med. Biol. **56**(14), 4399–4413 (2011).

24 C.B. Simone, D. Ly, T.D. Dan, J. Ondos, H. Ning, A. Belard, J. O’Connell, R.W. Miller, and N.L. Simone, “Comparison of intensity-modulated radiotherapy, adaptive radiotherapy, proton radiotherapy, and adaptive proton radiotherapy for treatment of locally advanced head and neck cancer.,” Radiother. Oncol. **101**(3), 376–82 (2011).

25 A.C. Kraan, S. van de Water, D.N. Teguh, A. Al-Mamgani, T. Madden, H.M. Kooy, B.J.M. Heijmen, and M.S. Hoogeman, “Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors.,” Int. J. Radiat. Oncol. Biol. Phys. **87**(5), 888–96 (2013).

26 W. Liu, X. Zhang, Y. Li, and R. Mohan, “Robust optimization of intensity modulated proton therapy.,” Med. Phys. **39**(2), 1079–91 (2012).

27 W. Liu, S.J. Frank, X. Li, Y. Li, P.C. Park, L. Dong, X. Ronald Zhu, and R. Mohan, “Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers.,” Med. Phys. **40**(5), 051711 (2013).

28 W. Liu, S.J. Frank, X. Li, Y. Li, R.X. Zhu, and R. Mohan, “PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization.,” Med. Phys. **40**(2), 021709 (2013).

29 A. Fredriksson, A. Forsgren, and B. Hårdemark, “Minimax optimization for handling range and setup uncertainties in proton therapy.,” Med. Phys. **38**(3), 1672–1684 (2011).

30 A. Fredriksson, “A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization.,” Med. Phys. **39**(8), 5169–81 (2012).

31 M. Stuschke, A. Kaiser, J. Abu Jawad, C. Pöttgen, S. Levegrün, and J. Farr, “Multi-scenario based robust intensity-modulated proton therapy (IMPT) plans can account for set-up errors more effectively in terms of normal tissue sparing than planning target volume (PTV) based intensity-modulated photon plans in the head and neck regi,” Radiat. Oncol. **8**(1), 145 (2013).

32 J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel, “Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity,” Med. Phys. **41**(2), 021722 (2014).

33 J.P. Tol, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel, “Increasing the number of arcs improves head and neck volumetric modulated arc therapy plans.,” Acta Oncol. **54**(2), 283–7 (2015).

34 P. Doornaert, W.F.A.R. Verbakel, D.H.F. Rietveld, B.J. Slotman, and S. Senan, “Sparing the contralateral submandibular gland without compromising PTV coverage by using volumetric modulated arc therapy.,” Radiat. Oncol. **6**(1), 74 (2011).

35 T.S. Hong, W.A. Tomé, R.J. Chappell, P. Chinnaiyan, M.P. Mehta, and P.M. Harari, “The impact of daily

setup variations on head-and-neck intensity-modulated radiation therapy.," Int. J. Radiat. Oncol. Biol. Phys. **61**(3), 779–88 (2005).

36 L. Gilbeau, M. Octave-Prignot, T. Loncol, L. Renard, P. Scalliet, and V. Grégoire, "Comparison of setup accuracy of three different thermoplastic masks for the treatment of brain and head and neck tumors.," Radiother. Oncol. **58**(2), 155–62 (2001).

37 M.F. Moyers, M. Sardesai, S. Sun, and D.W. Miller, "Ion stopping powers and CT numbers.," Med. Dosim. **35**(3), 179–194 (2010).

38 P.C. Park, X.R. Zhu, A.K. Lee, N. Sahoo, A.D. Melancon, L. Zhang, and L. Dong, "A Beam-Specific Planning Target Volume (PTV) Design for Proton Therapy to Account for Setup and Range Uncertainties," Int. J. Radiat. Oncol. **82**(2), e329–e336 (2012).

39 J.A. Langendijk, P. Lambin, D. De Ruysscher, J. Widder, M. Bos, and M. Verheij, "Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach.," Radiother. Oncol. **107**(3), 267–73 (2013).

Chapter 12

Discussion

Discussion

The general theme of this thesis is automating and improving the radiotherapy treatment planning process for head and neck cancer (HNC). Treatment of this disease site commonly includes two large and irregularly shaped planning target volumes (PTVs) requiring different prescription doses, and a large number of organs-at-risk (OARs) that are often in close proximity to the PTVs. Treatment planning for HNC is therefore a complex and labor-intensive process, making it an ideal paradigm to study advanced planning techniques. Automated treatment planning solutions have the potential to provide high quality planning results more consistently and may also improve efficiency in the radiotherapy department.

Two automated treatment planning solutions were discussed in this thesis. Firstly, **Chapter 5** provided a technical account of the in-house developed automatic interactive optimizer (AIO), along with a preliminary investigation of volumetric modulated arc therapy (VMAT) plan quality using an initial set of ten HNC patients. A more in-depth clinical evaluation on AIO plan quality was provided in **Chapter 6** using seventy HNC patients. In addition, versatility of the AIO solution was demonstrated by including twenty locally advanced lung cancer patients in the analysis.

Secondly, **Chapter 7** discusses RapidPlan, a commercial knowledge-based planning (KBP) solution developed by Varian Medical Systems (Palo Alto, USA). KBP can predict achievable OAR dose-volume histograms (DVHs) for individual patients using a model that correlates the geometric and dosimetric features of previously created plans. These achievable DVHs can be used as input for the subsequently performed optimization process. Dosimetric outliers in the model library, such as plans that provide sub-optimal OAR sparing, may deteriorate the accuracy of the DVH predictions. **Chapter 8** therefore investigated the effect on resulting plan quality when varying numbers of dosimetric outliers were included in a 70-patient HNC model.

A potential downside of using AIO is the dependency of the plans on the priorities that are used for the individual structures during the interactive optimization process. Plans resulting from RapidPlan, on the other hand, are dependent on the quality of the model library, while sub-optimal results can also be obtained if the prospective patient' s geometric features are not similar to those in the model library. A future study should assess the relative advantages and disadvantages of different automated planning solutions including AIO and RapidPlan for individual patients.

Benchmarked against manually interactively optimized plans, the AIO and RapidPlan solutions generally resulted in improved OAR sparing, while achieving similar PTV dose coverage and homogeneity values. The importance of this is demonstrated in **Chapter 3**,

where the trade-off between PTV dose homogeneity and OAR sparing was investigated for VMAT HNC plans, and an exponential relationship was found between both parameters. This relationship shows that small differences in PTV dose coverage and homogeneity may substantially influence OAR doses. As investigated in **Chapter 2**, the use of additional VMAT arcs may simultaneously improve OAR sparing and PTV dose homogeneity values. It is important to consider that a RapidPlan model library that, for example, consists solely of dual-arc VMAT plans, may underestimate the amount of achievable OAR sparing for treatment plans that utilize additional arcs. Accurate KBP OAR dose predictions therefore rely on the time-consuming process of creating different models to suit a particular indication or treatment technique. This is of no concern with AIO because the level of OAR sparing that can be achieved by the inclusion of additional arcs will become apparent during the interactive optimization process, as AIO adjusts the placement of the optimization objectives accordingly.

Chapter 9 investigated the use of RapidPlan to provide patient-specific plan quality assurance (QA) by comparing the predicted OAR DVH-line with the DVH that was obtained during the treatment planning process. Composition of the model library is important if RapidPlan is used for QA, because accurate and realistic OAR DVH predictions require similar planning techniques to be used for the plans in the model library and for the plan that is to be evaluated. RapidPlan could also be used to provide plan QA of plans admitted to clinical trials. As investigated in **Chapter 4**, however, the mandated PTV dose coverage and homogeneity criteria can vary between the planning protocols of different trials, substantially influencing achievable levels of OAR sparing. Using RapidPlan for clinical trial plan QA therefore necessitates the creation of a model library with plans that provide PTV dose coverage and homogeneity values considered acceptable by the clinical trial. **Chapter 4** emonstrated that using the PTV dose coverage and homogeneity criteria specified by the European Organization for Research and the Treatment of Cancer (EORTC) instead of those of the Radiation Therapy and Oncology Group (RTOG) could, on average, improve mean dose values to the salivary glands and swallowing muscles by 7.6Gy and 7.0Gy, respectively. It is interesting to note that these gains are of the same magnitude as those reported when comparing intensity modulated proton therapy and VMAT, as investigated in **Chapter 11**.

The AIO solution has been in routine clinical use to automate the interactive optimization process for clinical HNC plans at the department of radiation oncology of the VU University Medical Center since February 2014. The initial studies presented in **Chapters 5** and **6** showed that AIO optimized plans were at least of similar quality as their manually optimized

counterparts. These studies, however, were performed in a controlled environment using the same set of patients to benchmark both techniques, making it uncertain whether the predicted gains have actually been translated in routine clinical practice. Therefore, to verify that novel dose delivery and treatment planning techniques have been successfully introduced in clinical practice, **Chapter 10** provided a longitudinal evaluation of routine clinical HNC plans delivered at our department in the past decade. Comparing the first period (manually interactively optimized intensity modulated radiotherapy [IMRT] plans from 2005-2008), with the second period (covering the introduction of VMAT for HNC), third period (aiming for further improvements in OAR sparing while including the submandibular glands and swallowing muscles) and fourth period (AIO optimized dual-arc VMAT plans from 2014-2015) showed that substantial gains in OAR sparing were achieved over the periods. These gains did not degrade PTV dose homogeneity, which is important considering the aforementioned trade-off with OAR sparing, nor did they increase dose deposition in the remainder of the body. While the use of more than two arcs for HNC planning has the potential to further improve plan quality (**Chapter 2**), this has yet to be more widely implemented in clinical practice. Although four arc VMAT plans are used occasionally at the department of radiation oncology of the VU University Medical Center for complex HNC patients, it remains difficult to predict which patients benefit the most from using more than two arcs. A potential downside is the increased delivery time compared to dual-arc plans.

The improvements in clinical HNC plans noted in **Chapter 10** are important when considering the potential future role for proton therapy in the treatment of HNC patients. While planning comparisons between proton and photon plans should identify individual patients that may benefit from proton therapy, it is important that optimal photon plans are considered for this comparison. Furthermore, as investigated in **Chapter 11**, it is important to simultaneously benchmark OAR sparing, PTV dose homogeneity and plan robustness in the comparison.

IMRT and VMAT Optimization

The clinical use of IMRT, or its rotational variant VMAT, has increased in recent years[1,2]. Since these techniques have been demonstrated to improve plan quality, i.e. improving OAR sparing while achieving similar PTV dose coverage and homogeneity values, for a large number of disease sites[3,4], they have largely replaced three-dimensional conformal radiotherapy (3D-CRT) as the standard of care in modern day radiotherapy. The optimization phase for IMRT planning may be approached in various ways. Some institutes create

templates and use the same set of optimization objectives to create plans for all patients of a particular disease site[5,6]. Although such templates may result in acceptable plan quality for the majority of patients, the nuances of a specific patient's geometry may prevent optimal plans from being obtained in individual patients[7]. This problem may be solved by performing subsequent additional optimizations while adjusting the optimization objectives, thereby iteratively navigating towards optimal plans for individual patients[8]. This approach, however, is labor-intensive, time consuming, and subject to variation[9–11]. Nelms et al.[11] performed a detailed comparison of external beam treatment planning quality between different planners, institutes and treatment planning systems (TPSs) using a plan quality metric (PQM). They concluded that there was large variation between the planners' ability to meet the various plan objectives quantified by the PQM, while no statistically differences were found obtained between different TPSs, delivery techniques and plan complexity. Interestingly, certification, education, experience levels and confidence levels were not found to be good predictors of resulting plan quality. Das et al.[12] performed a retrospective analysis of brain, head and neck and prostate IMRT plans, and found substantial variation in the prescribed and delivered doses between institutes. Given these differences, concerns were raised about the validity of comparing clinical outcomes for IMRT between different institutes.

Automated treatment planning

Automated solutions to treatment planning typically remove or reduce the manual interaction between the user and planning system, with the aim of creating high quality plans more consistently[13]. Two distinct approaches have received most attention in the literature: knowledge-based planning (KBP) and multicriteria optimization (MCO).

Knowledge-based treatment planning

As described previously, KBP involves the creation of a model derived from the geometric and dosimetric features of previously created treatment plans[14–17]. If the geometry of a new patient is similar to the plans that were used to create the model, KBP can be used to predict achievable OAR doses[18], which may in turn be used to estimate optimal placement of the optimization objectives[19–21] and associated priorities[22–24].

Based on initial work by the Washington University School of Medicine[14] and the Duke University Medical Center[15], Varian Medical Systems recently released RapidPlan as a commercial KBP solution. Fogliata et al.[25,26] used RapidPlan to automate VMAT planning for prostate, lung and hepatocellular cancer patients. For all disease sites, the quality of the plans resulting from RapidPlan were found to be comparable to the manually optimized

plans that were used clinically, leading the authors to conclude that RapidPlan might be considered as a tool to streamline and improve the planning process. Comparable results were noted for locally advanced HNC patients in **Chapter 7,** although deviations could be noted if the patient was geometrically different from the bulk of the model.

Automated multicriteria optimization

While many ways to perform automated MCO are described in the literature, the approach generally consists of automatically performing consecutive optimization runs and navigating towards optimal placement of the objectives for individual patients. By iteratively changing the optimization objectives that are used, dose distributions that better satisfy the plan evaluation criteria are created during each optimization run. The working of AIO (**Chapters 5** and **6**) is similar to an automated MCO solution but, instead of performing many subsequent optimization runs, it utilizes the interactive optimization process allowed by the Eclipse TPS to iteratively adjust the optimization objectives attempting to maximize OAR sparing throughout a single optimization.

iCycle[27], an in-house developed automatic MCO solution from the Erasmus University Medical Center, uses a wish-list that specifies both hard constraints and planning priorities to guide the optimization runs[28,29]. The hard constraints indicate optimization objectives that must be satisfied, and for example include maximum doses to serial OARs, and minimum PTV dose requirements. The ordering of the various planning priorities in the wish-list indicates the relative importance in satisfying the specified goals. By also incorporating beam angle optimization[30], the iCycle solution can automatically generate non-coplanar IMRT plans[31]. This has been applied successfully to HNC IMRT plan optimization[32]. Depending on the number of IMRT beams, generating coplanar plans (including optimization) using iCycle could take 1-2 hours[27]. Recent studies have also used iCycle for prostate and cervical cancer VMAT plans[33,34], and automatic adaptive replanning was demonstrated for liver stereotactic body radiotherapy[35]. It should be noted that iCycle does not perform automated MCO for VMAT plans and instead uses the results of IMRT MCO as input for the VMAT optimization[36].

The concept of automated MCO has been incorporated by Philips Radiation Oncology Systems as Auto-Planning in the Pinnacle TPS, which uses an iterative algorithm based approach to automatically adapt objectives, constraints and dose shaping contours during the optimization process to achieve clinically set goals. Krayenbuehl et al.[37] evaluated Pinnacle Auto-Planning using fifty HNC VMAT patients and, compared to manually created plans, found significant reductions in mean doses to the parotid glands, oral mucosa, swallowing muscles and dorsal neck tissue, along with a reduced maximum dose

to the spinal cord. The effective working time, i.e. the time that the planner had to interact with the TPS, was shortened from approximately 48.5 to 3.8 minutes, although they did not report the computation times. While the effective working time was not investigated for RapidPlan or AIO, it is likely to be of the same order of magnitude.

Automated MCO has been used by Craft et al. to create large numbers of Pareto-optimal plans[38]. These were used to span the Pareto-surface, consisting of plans with different trade-offs between the various planning goals[39]. The user can navigate the Pareto-surface and select a plan with the desired clinical trade-offs. For IMRT plans of the pancreas, head and neck and glioblastoma, this technique has been shown to decrease the time required for the treatment planning process while improving plan quality[40,41]. Creation of VMAT plans was also shown to be possible by, for example, using the IMRT plan DVHs as input optimization objectives for VMAT[36,42].

Pareto-front navigation has been implemented in the Raystation TPS (Raysearch laboratories, Sweden) based on the aforementioned investigations[43]. In this system, the user can adjust sliders denoting the relative importance of the different optimization structures to select a clinical plan. Plans created using this technique were benchmarked against manually created plans in a recent investigation by Kierkels et al.[44], which included a blinded assessment of the plans. Plans were found to be highly similar, while the time required to create a MCO plan was 43 minutes, compared to 205 minutes for the manually created plans.

Although the large number of optimizations that are often required to create the Pareto-front can make it more time consuming than using KBP to generate a plan, it should be noted that the creation of a KBP model can take a substantial amount of time as well. Additionally, since MCO does not require the creation of a model, it is expected to work irrespective of the patient's geometrical features. Time spent to create a KBP model, however, is a one-time investment after which optimal optimization objective placement can be determined within a minute and the time requirements for the optimization and dose calculations are typically less than twenty minutes.

While automated treatment planning solutions are relatively new, and few have been introduced in routine clinical practice, automatic navigation towards high-quality patient-specific treatment plans follows the current trend of personalized medicine. Although plans created using the automated methods generally improved quality over their manually created counterparts, it should be noted that the absolute quality of these plans may vary substantially, as demonstrated by the work in this thesis. Future work is therefore warranted

to benchmark different automated planning solutions. It should also be noted that many of aforementioned automated planning solutions revolve around removing the variability of the user during the optimization process. Fully automating the treatment planning process, including the automatic segmentation of tumor and organ-at-risk volumes and positioning of treatment fields, requires a clear understanding of the relationship between dosimetric factors and OAR toxicity and necessitates further work.

In addition, graphics processing unit (GPU)-based programming has the potential to substantially decrease time requirements for computerized algorithms[45,46]. Compared to regular central processing units (CPUs), GPUs typically have far more processing cores, particularly allowing parallel computer tasks to be performed faster. The advent of GPU-based programming in radiotherapy will in the near future likely allow optimizations with full heterogeneity correction and subsequent dose calculations to be performed in near real time, further improving the time efficiency of automated planning solutions.

Evaluation of plan quality

Objectively evaluating the quality of a given treatment plan may be difficult for various reasons, including i) depending on their unique geometric characteristics, individual patients may present different trade-offs between the varying planning objectives, ii) patient-specific levels of achievable OAR sparing are often unknown before the treatment planning process is started, and iii) interpreting the clinical implications of OAR dose reductions is not straightforward, especially when many individual OARs are included in a given plan.

Using their in-house developed KBP solution, Moore et al.[47] recently quantified unnecessary normal tissue complication risks due to suboptimal planning for prostate cancer in the RTOG 0126 study. They identified that 43% patients had a ≥5% excess risk for rectal toxicity as a result of sub-optimal planning. Stanhope et al.[48] used a similar approach, also with an in-house solution, to evaluate the effect of various simulated delivery errors on the resulting OAR DVHs. These studies performed retrospective analyses to identify the effect of either sub-optimal planning or delivery errors on resulting OAR doses and subsequent normal tissue complication probability (NTCP) values. As suggested in **Chapter 9**, using a KBP model to QA a plan before admitting it to a clinical trial provides an accurate and fast benchmark, and may prevent sub-optimal plans from being included in the clinical trial. This idea has yet to be incorporated in clinical trials.

Future Perspectives

A large number of advances have been introduced in modern radiotherapy over the past decade. The expected effect of novel dose delivery (e.g., using IMRT and VMAT instead of 3D-CRT) or treatment planning techniques (e.g., evaluating the effect of including additional OARs for sparing) are often based on 'traditional' treatment planning studies, aiming to benchmark these different techniques on the same set of patients. While such studies may objectively assess whether treatment plans improve using new techniques, it remains uncertain whether the predicted gains are successfully being translated into routine clinical practice. **Chapter 10** therefore evaluated a large number of clinical treatment plans for HNC that were delivered at our department in the past decade. We think that further improvements in radiotherapy using photons are still feasible. Further understanding of the relationship between dose and OAR toxicity will facilitate the translation of different clinical goals into planning objectives and help improve future plans. Future studies should ideally also correlate the dosimetric gains with patient-reported outcome measures (PROMs) to investigate whether improved OAR sparing results in a decrease in long-term adverse effects of radiotherapy and improves the quality of life[49].

Intensity modulated proton therapy (IMPT), has been predicted by some to benefit a large proportion of HNC patients. In theory, protons have the potential to further improve OAR sparing because their finite range causes a steep dose fall-off after its dose maximum (the so-called 'Bragg-peak"). This phenomenon can be used to minimize doses to structures lying behind the target in the beam's eye view of the individual fields. However, uncertainties in patient positioning (setup errors) and Hounsfield density values of the planning CT-scan (range uncertainties) should be taken into account during treatment planning for proton therapy because they may degrade the apparently achievable levels of OAR sparing and because such errors can influence the location of the Bragg-peak leading to either underdosage of the target, or a higher OAR dose[50]. As a result, the PTV concept, introduced in photon radiotherapy to account for uncertainties in patient positioning, is no longer valid in proton therapy, instead requiring robust optimization methods in which the uncertainties are taken into account during the optimization process.

As demonstrated in **Chapter 11**, trade-offs between plan robustness and OAR sparing were found between different proton therapy techniques. Using 3 treatment fields, significant improvements in plan robustness were found using single-field optimization (SFO) rather than multi-field optimization (MFO), although swallowing muscle and salivary gland mean doses were respectively 6.9Gy and 5.4Gy higher. This was also noted by Quan et al.[51], who showed

that 3-field MFO plans improved parotid gland mean doses by 6.7Gy on average, compared to 3-field SFO plans. The authors, however, also noted an average mean dose increase of 9.9Gy to the oral cavity in the MFO plans, while this structure was not evaluated in our results. The field orientations may have also contributed to these results. In **Chapter 11** it was also noted that 7-field MFO plans improved mean doses to the swallowing muscles and salivary glands by 8.9Gy and 9.5Gy, respectively, compared to their respective VMAT plans, although this came at the expense of the poorest plan robustness; Modeling the worst case scenario, which combined ±3mm setup errors (in three directions) with ±3% range uncertainties, resulted in a boost clinical target volume (CTV_B) V95% of only 79.9% for one patient, while this was 87.9% for the 3-field SFO plan. This was consistent with Albertini et al.[52], who showed that SFO plans were more robust than MFO plans for both set-up and range errors.

Robust optimization techniques for IMPT may incorporate potential setup errors and range uncertainties during the optimization process[53]. Such techniques allow the user to simultaneously minimize OAR doses while still ensuring adequate CTV dose coverage values under the combined influence of the range and set-up uncertainties that have been modeled. Since appropriate margins are determined during the robust optimization process, this avoids the need for using uniform (conventional) margins during planning for IMPT, which could reduce the size of the irradiated target volume and may improve OAR sparing. In a recent study by Liu et al.[54], robust optimized IMPT plans (using CTVs as the target) were benchmarked against their non-robustly optimized counterparts (using conventional PTVs). Averaged over 14 patients they found that, in their worst-case scenario that modeled ±3mm setup errors and ±3.5% range uncertainties, robustly optimized plans had improved target dose coverage over their non-robust counterparts, with a 2.7% dose increase to 95% of the gross tumor volume (GTV D95%), and a 2.0% increase to primary CTV D95%. Robust optimization also lead to modest improvements in OAR sparing, with mean dose improvements to the oral cavity, left and right parotid glands of 0.7Gy, 1.4Gy and 0.7Gy, respectively. Similar conclusions were reached by Fredriksson et al.[53] for a lung case, a paraspinal case, and a prostate case.

MRI-guided radiotherapy (MRIgRT) is another recent advance offering the potential of improving radiotherapy treatments by combining radiation dose delivery with magnetic resonance imaging (MRI). MRIgRT has been approached in multiple ways; The University Medical Center Utrecht pioneered the integration of a 6MV linear accelerator with 1.5T MRI imaging in the MR-linac, in collaboration with Elekta AB (Stockholm, Sweden), and Philips (Best, The Netherlands)[55]. ViewRay Inc. (Oakwood, OH, USA) has recently developed the MRIdian, combining 0.35T MR-imaging with three cobalt60 (^{60}Co) sources combined with multileaf collimators (MLCs) to permit multiple-beam IMRT[56]. Both systems also allow for images to

be obtained in the treatment position prior to irradiation, as well as during irradiation. This should aid the verification of radiation delivery / targeting and, when combined with deformable image registration and fast dose optimization / calculation algorithms, should facilitate adaptive replanning techniques that take daily anatomical geometric changes into account[57].

Adaptive radiotherapy has the potential to improve treatment for HNC patients in multiple ways, including: 1) Salivary glands have been shown to both reduce in volume, and move inwards over the course of radiotherapy. Barker et al.[58] found a median shift for the parotid glands of 3.1mm (range: 0-9.9mm) over the course of radiotherapy, along with a daily median volume reduction of 0.19cm^3. Adaptive replanning may account for such anatomical changes and provide improved sparing of the salivary glands throughout the fractions[59]. 2) HNC patients often present large tumors that may shrink substantially over the course of radiotherapy. Adaptive replanning techniques may account for the volume changes of these tumors and, since the total target volume becomes smaller over the course of the treatments, this may contribute to improved OAR sparing. Hansen et al.[60] found that adaptive replanning was necessary to ensure adequate target coverage throughout the course of radiotherapy given these anatomical changes. Up to date however, there is no consensus about reducing PTV margins during the course of treatment based on the reductions of the visible gross tumor volume. While the tumor may appear smaller on the CT and MRI scans, this does not exclude the presence of microscopic tumor cells in the regressed tumor volume. Clinical trials are therefore necessary to evaluate whether adaptive planning strategies can be performed based on the reduction in visible tumor size.

It is important to note that although the various technical advances described in this thesis have the potential to further improve radiotherapy treatment outcomes, accurate delineation of tumor regions and organs-at-risk remains an integral part of providing effective radiotherapy treatments. Target volume delineation has been shown to vary greatly amongst observers[61], despite the use of positron emission tomography (PET)-CT in some locations[62,63]. Similarly, Nelms et al.[64] has found significant inter-clinician variation in OAR contouring in the head and neck region, resulting in mean dose differences ranging from -289% to 56%. Simultaneous improvements in target and OAR delineation will be needed to fully realize the potential of technical advances in radiotherapy.

This thesis highlighted that substantial improvements in IMRT and VMAT treatment plan quality have been made over the years. Up to date, these gains have been reserved solely for modern radiotherapy institutes, with sufficient staff to explore the optimal plans for individual patients. The advent of automated planning, however, may contribute to the widespread creation of such plans in radiotherapy institutes that have less resources to spare.

References

1 D.A. Shumway, K.A. Griffith, L.J. Pierce, M. Feng, J.M. Moran, M.H. Stenmark, R. Jagsi, and J.A. Hayman, "Wide Variation in the Diffusion of a New Technology: Practice-Based Trends in Intensity-Modulated Radiation Therapy (IMRT) Use in the State of Michigan, With Implications for IMRT Use Nationally.," J. Oncol. Pract. 11(3), e373–9 (2015).

2 K. Bak, E. Murray, E. Gutierrez, J. Ross, and P. Warde, "IMRT utilization in Ontario: qualitative deployment evaluation.," Int. J. Health Care Qual. Assur. **27**(8), 742–59 (2014).

3 A.M. Chen, B.Q. Li, D.G. Farwell, J. Marsano, S. Vijayakumar, and J.A. Purdy, "Improved dosimetric and clinical outcomes with intensity-modulated radiotherapy for head-and-neck cancer of unknown primary origin.," Int. J. Radiat. Oncol. Biol. Phys. **79**(3), 756–62 (2011).

4 J.M. Michalski, Y. Yan, D. Watkins-Bruner, W.R. Bosch, K. Winter, J.M. Galvin, J.P. Bahary, G.C. Morton, M.B. Parliament, and H.M. Sandler, "Preliminary toxicity analysis of 3-dimensional conformal radiation therapy versus intensity modulated radiation therapy on the high-dose arm of the Radiation Therapy Oncology Group 0126 prostate cancer trial.," Int. J. Radiat. Oncol. Biol. Phys. **87**(5), 932–8 (2013).

5 D. Georg, B. Kroupa, P. Georg, P. Winkler, J. Bogner, K. Dieckmann, and R. Pötter, "Inverse planning--a comparative intersystem and interpatient constraint study.," Strahlenther. Onkol. **182**(8), 473–80 (2006).

6 P. Xia, N. Lee, Y.M. Liu, I. Poon, V. Weinberg, E. Shin, J.M. Quivey, and L.J. Verhey, "A study of planning dose constraints for treatment of nasopharyngeal carcinoma using a commercial inverse treatment planning system.," Int. J. Radiat. Oncol. Biol. Phys. **59**(3), 886–96 (2004).

7 H.W. Hamacher and K.H. Küfer, "Inverse radiation therapy planning — a multiple objective optimization approach," Discret. Appl. Math. **118**(1-2), 145–161 (2002).

8 A.J. Mundt and J.C. Roeske, "Intensity Modulated Radiation Therapy: *A Clinical Perspective*.,". BC Decker Inc., Hamilton, London (2005).

9 J.J. Wilkens, J.R. Alaly, K. Zakarian, W.L. Thorstad, and J.O. Deasy, "IMRT treatment planning based on prioritizing prescription goals.," Phys. Med. Biol. **52**(6), 1675–92 (2007).

10 Y. Matsuo, K. Takayama, Y. Nagata, E. Kunieda, K. Tateoka, N. Ishizuka, T. Mizowaki, Y. Norihisa, M. Sakamoto, Y. Narita, S. Ishikura, and M. Hiraoka, "Interinstitutional variations in planning for stereotactic body radiation therapy for lung cancer.," Int. J. Radiat. Oncol. Biol. Phys. **68**(2), 416–25 (2007).

11 B.E. Nelms, G. Robinson, J. Markham, K. Velasco, S. Boyd, S. Narayan, J. Wheeler, and M.L. Sobczak, "Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems," Pract. Radiat. Oncol. **2**(4), 296–305 (2012).

12 I.J. Das, C.W. Cheng, K.L. Chopra, R.K. Mitra, S.P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems.," J. Natl. Cancer Inst. **100**(5), 300–7 (2008).

13 M.B. Sharpe, K.L. Moore, and C.G. Orton, "Point/Counterpoint: Within the next ten years treatment planning will become fully automated without the need for human intervention.," Med. Phys. **41**(12), 120601 (2014).

14 L.M. Appenzoller, J.M. Michalski, W.L. Thorstad, S. Mutic, and K.L. Moore, "Predicting dose-volume histograms for organs-at-risk in IMRT planning.," Med. Phys. **39**(12), 7446–61 (2012).

15 L. Yuan, Y. Ge, W.R. Lee, F.F. Yin, J.P. Kirkpatrick, and Q.J. Wu, "Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans.," Med. Phys. **39**(11), 6868–78 (2012).

16 B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, M. Chuang, R. Taylor, R. Jacques, and T. McNutt, "Patient geometry-driven information retrieval for IMRT treatment plan quality control.," Med. Phys. **36**(12), 5497–505 (2009).

17 J. Lian, L. Yuan, Y. Ge, B.S. Chera, D.P. Yoo, S. Chang, F. Yin, and Q.J. Wu, "Modeling the dosimetry of organ-at-risk in head and neck IMRT planning: An intertechnique and interinstitutional study.," Med. Phys. **40**(12), 121704 (2013).

[18] X. Zhu, Y. Ge, T. Li, D. Thongphiew, F.F. Yin, and Q.J. Wu, "A planning quality evaluation tool for prostate adaptive IMRT based on machine learning," Med. Phys. **38**(2), 719 (2011).

[19] Y. Yang, E.C. Ford, B. Wu, M. Pinkawa, B. van Triest, P. Campbell, D.Y. Song, and T.R. McNutt, "An overlap-volume-histogram based method for rectal dose prediction and automated treatment planning in the external beam prostate radiotherapy following hydrogel injection.," Med. Phys. **40**(1), 011709 (2013).

[20] B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning.," Int. J. Radiat. Oncol. Biol. Phys. **79**(4), 1241–7 (2011).

[21] D. Good, J. Lo, W.R. Lee, Q.J. Wu, F.F. Yin, and S.K. Das, "A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning.," Int. J. Radiat. Oncol. Biol. Phys. **87**(1), 176–81 (2013).

[22] J.J. Boutilier, T. Lee, T. Craig, M.B. Sharpe, and T.C.Y. Chan, "Models for predicting objective function weights in prostate cancer IMRT.," Med. Phys. **42**(4), 1586–95 (2015).

[23] T. Lee, M. Hammad, T.C.Y. Chan, T. Craig, and M.B. Sharpe, "Predicting objective function weights from patient anatomy in prostate IMRT treatment planning.," Med. Phys. **40**(12), 121706 (2013).

[24] M. Zarepisheh, T. Long, N. Li, Z. Tian, H.E. Romeijn, X. Jia, and S.B. Jiang, "A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning.," Med. Phys. **41**(6), 061711 (2014).

[25] A. Fogliata, F. Belosi, A. Clivio, P. Navarria, G. Nicolini, M. Scorsetti, E. Vanetti, and L. Cozzi, "On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer.," Radiother. Oncol. **113**(3), 385–91 (2014).

[26] A. Fogliata, P.M. Wang, F. Belosi, A. Clivio, G. Nicolini, E. Vanetti, and L. Cozzi, "Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer.," Radiat. Oncol. **9**(1), 236 (2014).

[27] S. Breedveld, P.R.M. Storchi, P.W.J. Voet, and B.J.M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans.," Med. Phys. **39**(2), 951–63 (2012).

[28] S. Breedveld, P.R.M. Storchi, M. Keijzer, A.W. Heemink, and B.J.M. Heijmen, "A novel approach to multi-criteria inverse planning for IMRT.," Phys. Med. Biol. **52**(20), 6339–53 (2007).

[29] S. Breedveld, P.R.M. Storchi, and B.J.M. Heijmen, "The equivalence of multi-criteria methods for radiotherapy plan optimization.," Phys. Med. Biol. **54**(23), 7199–209 (2009).

[30] S. Das, T. Cullip, G. Tracton, S. Chang, L. Marks, M. Anscher, and J. Rosenman, "Beam orientation selection for intensity-modulated radiation therapy based on target equivalent uniform dose maximization.," Int. J. Radiat. Oncol. Biol. Phys. **55**(1), 215–24 (2003).

[31] P.W.J. Voet, S. Breedveld, M.L.P. Dirkx, P.C. Levendag, and B.J.M. Heijmen, "Integrated multicriterial optimization of beam angles and intensity profiles for coplanar and noncoplanar head and neck IMRT and implications for VMAT.," Med. Phys. **39**(8), 4858–65 (2012).

[32] P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, D. Fransen, P.C. Levendag, and B.J.M. Heijmen, "Toward fully automated multicriterial plan generation: a prospective clinical study.," Int. J. Radiat. Oncol. Biol. Phys. **85**(3), 866–72 (2013).

[33] A.W.M. Sharfo, P.W.J. Voet, S. Breedveld, J.W.M. Mens, M.S. Hoogeman, and B.J.M. Heijmen, "Comparison of VMAT and IMRT strategies for cervical cancer patients using automated planning," Radiother. Oncol. (2015).

[34] P.W.J. Voet, M.L.P. Dirkx, S. Breedveld, A. Al-Mamgani, L. Incrocci, and B.J.M. Heijmen, "Fully automated volumetric modulated arc therapy plan generation for prostate cancer patients.," Int. J. Radiat. Oncol. Biol. Phys. **88**(5), 1175–9 (2014).

[35] S.M. Leinders, S. Breedveld, A. Méndez Romero, D. Schaart, Y. Seppenwoolde, and B.J.M. Heijmen,

"Adaptive liver stereotactic body radiation therapy: automated daily plan reoptimization prevents dose delivery degradation caused by anatomy deformations.," Int. J. Radiat. Oncol. Biol. Phys. **87**(5), 1016–21 (2013).

[36] H. Chen, D.L. Craft, and D.P. Gierga, "Multicriteria optimization informed VMAT planning.," Med. Dosim. **39**(1), 64–73 (2014).

[37] J. Krayenbuehl, I. Norton, G. Studer, and M. Guckenberger, "Evaluation of an automated knowledge based treatment planning system for head and neck.," Radiat. Oncol. **10**(1), 226 (2015).

[38] D.L. Craft, T.F. Halabi, H.A. Shih, and T.R. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning.," Med. Phys. **33**(9), 3399–407 (2006).

[39] D. Craft and C. Richter, "Deliverable navigation for multicriteria step and shoot IMRT treatment planning.," Phys. Med. Biol. **58**(1), 87–103 (2013).

[40] D.L. Craft, T.S. Hong, H.A. Shih, and T.R. Bortfeld, "Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy.," Int. J. Radiat. Oncol. Biol. Phys. **82**(1), e83–90 (2012).

[41] T.S. Hong, D.L. Craft, F. Carlsson, and T.R. Bortfeld, "Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head.," Int. J. Radiat. Oncol. Biol. Phys. **72**(4), 1208–14 (2008).

[42] D. Craft, D. McQuaid, J. Wala, W. Chen, E. Salari, and T. Bortfeld, "Multicriteria VMAT optimization.," Med. Phys. **39**(2), 686–96 (2012).

[43] Multi-criteria optimization in Raystation. Stockhol: RaySearch Laboratories AB (2012).

[44] R.G.J. Kierkels, R. Visser, H.P. Bijl, J.A. Langendijk, A.A. van 't Veld, R.J.H.M. Steenbakkers, and E.W. Korevaar, "Multicriteria optimization enables less experienced planners to efficiently produce high quality treatment plans in head and neck cancer radiotherapy.," Radiat. Oncol. **10**(1), 87 (2015).

[45] X. Jia, X. Gu, Y.J. Graves, M. Folkerts, and S.B. Jiang, "GPU-based fast Monte Carlo simulation for radiotherapy dose calculation.," Phys. Med. Biol. **56**(22), 7017–31 (2011).

[46] C. Men, X. Jia, and S.B. Jiang, "GPU-based ultra-fast direct aperture optimization for online adaptive radiation therapy.," Phys. Med. Biol. **55**(15), 4309–19 (2010).

[47] K.L. Moore, R. Schmidt, V. Moiseenko, L.A. Olsen, J. Tan, Y. Xiao, J. Galvin, S. Pugh, M.J. Seider, A.P. Dicker, W. Bosch, J. Michalski, and S. Mutic, "Quantifying Unnecessary Normal Tissue Complication Risks due to Suboptimal Planning: A Secondary Study of RTOG 0126.," Int. J. Radiat. Oncol. Biol. Phys. **92**(2), 228–35 (2015).

[48] C. Stanhope, Q.J. Wu, L. Yuan, J. Liu, R. Hood, F.F. Yin, and J. Adamson, "Utilizing knowledge from prior plans in the evaluation of quality assurance.," Phys. Med. Biol. **60**(12), 4873–91 (2015).

[49] V. Grégoire, J.A. Langendijk, and S. Nuyts, "Advances in Radiotherapy for Head and Neck Cancer.," J. Clin. Oncol. **33**(29), 3277–84 (2015).

[50] A.C. Kraan, S. van de Water, D.N. Teguh, A. Al-Mamgani, T. Madden, H.M. Kooy, B.J.M. Heijmen, and M.S. Hoogeman, "Dose uncertainties in IMPT for oropharyngeal cancer in the presence of anatomical, range, and setup errors.," Int. J. Radiat. Oncol. Biol. Phys. **87**(5), 888–96 (2013).

[51] E.M. Quan, W. Liu, R. Wu, Y. Li, S.J. Frank, X. Zhang, X.R. Zhu, and R. Mohan, "Preliminary evaluation of multifield and single-field optimization for the treatment planning of spot-scanning proton therapy of head and neck cancer.," Med. Phys. **40**(8), 081709 (2013).

[52] F. Albertini, E.B. Hug, and A.J. Lomax, "Is it necessary to plan with safety margins for actively scanned proton therapy?," Phys. Med. Biol. **56**(14), 4399–4413 (2011).

[53] A. Fredriksson, A. Forsgren, and B. Hårdemark, "Minimax optimization for handling range and setup uncertainties in proton therapy.," Med. Phys. **38**(3), 1672–1684 (2011).

[54] W. Liu, S.J. Frank, X. Li, Y. Li, P.C. Park, L. Dong, X. Ronald Zhu, and R. Mohan, "Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers.," Med. Phys. **40**(5), 051711 (2013).

[55] J.J.W. Lagendijk, B.W. Raaymakers, and M. van Vulpen, "The magnetic resonance imaging-linac sys-

tem.," Semin. Radiat. Oncol. **24**(3), 207–9 (2014).

[56] S. Mutic and J.F. Dempsey, "The ViewRay system: magnetic resonance-guided and controlled radiotherapy.," Semin. Radiat. Oncol. **24**(3), 196–9 (2014).

[57] S. Acharya, B.W. Fischer-Valuck, R. Kashani, P. Parikh, D. Yang, T. Zhao, O. Green, O. Wooten, H.H. Li, Y. Hu, V. Rodriguez, L. Olsen, C. Robinson, J. Michalski, S. Mutic, and J. Olsen, "Online Magnetic Resonance Image Guided Adaptive Radiation Therapy: First Clinical Applications.," Int. J. Radiat. Oncol. Biol. Phys. **94**(2), 394–403 (2015).

[58] J.L. Barker, A.S. Garden, K.K. Ang, J.C. O'Daniel, H. Wang, L.E. Court, W.H. Morrison, D.I. Rosenthal, K.S.C. Chao, S.L. Tucker, R. Mohan, and L. Dong, "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system.," Int. J. Radiat. Oncol. Biol. Phys. **59**(4), 960–70 (2004).

[59] J. Castelli, A. Simon, G. Louvel, O. Henry, E. Chajon, M. Nassef, P. Haigron, G. Cazoulat, J.D. Ospina, F. Jegoux, K. Benezery, and R. de Crevoisier, "Impact of head and neck cancer adaptive radiotherapy to spare the parotid glands and decrease the risk of xerostomia.," Radiat. Oncol. **10**(1), 6 (2015).

[60] E.K. Hansen, M.K. Bucci, J.M. Quivey, V. Weinberg, and P. Xia, "Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer.," Int. J. Radiat. Oncol. Biol. Phys. **64**(2), 355–62 (2006).

[61] R.J.H.M. Steenbakkers, J.C. Duppen, I. Fitton, K.E.I. Deurloo, L. Zijp, A.L.J. Uitterhoeve, P.T.R. Rodrigus, G.W.P. Kramer, J. Bussink, K. De Jaeger, J.S.A. Belderbos, A.A.M. Hart, P.J.C.M. Nowak, M. van Herk, and C.R.N. Rasch, "Observer variation in target volume delineation of lung cancer related to radiation oncologist-computer interaction: a 'Big Brother' evaluation.," Radiother. Oncol. **77**(2), 182–90 (2005).

[62] R.J.H.M. Steenbakkers, J.C. Duppen, I. Fitton, K.E.I. Deurloo, L.J. Zijp, E.F.I. Comans, A.L.J. Uitterhoeve, P.T.R. Rodrigus, G.W.P. Kramer, J. Bussink, K. De Jaeger, J.S.A. Belderbos, P.J.C.M. Nowak, M. van Herk, and C.R.N. Rasch, "Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis.," Int. J. Radiat. Oncol. Biol. Phys. **64**(2), 435–48 (2006).

[63] C.M. Anderson, W. Sun, J.M. Buatti, J.E. Maley, B. Policeni, S.L. Mott, and J.E. Bayouth, "Interobserver and intermodality variability in GTV delineation on simulation CT, FDG-PET, and MR Images of Head and Neck Cancer.," Jacobs J. Radiat. Oncol. **1**(1), 006 (2014).

[64] B.E. Nelms, W.A. Tomé, G. Robinson, and J. Wheeler, "Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer.," Int. J. Radiat. Oncol. Biol. Phys. **82**(1), 368–78 (2012).

Summary

The three major treatment options for cancer are generally recognized as being surgery, radiotherapy and chemotherapy. Many factors determine which treatment, or combination thereof, is chosen for a specific patient, including the site and stage of the disease, and the patient's biological age, overall health and personal preference. Radiotherapy plays an important role in the treatment of locally advanced head and neck cancer (HNC) and, when combined with chemotherapy, at least as effective as surgical resection for the majority of patients, with the possibility of organ-preservation as an added advantage.

Radiotherapy treatments require the creation of a treatment plan that should result in the delivery of sufficient dose to the tumor, whilst ideally sparing the normal tissues as much as possible. In current clinical practice, treatment plans are often made using computer algorithms with the final plan being obtained through an iterative process in which the resulting improvements get progressively smaller. The quality of the resulting plan, however, strongly depends on the skill of the planner and choices made by the treating physician, which can result in substantial differences between the plans from different planners and radiotherapy departments. This thesis therefore revolves around improving and automating the radiotherapy treatment planning process for HNC.

Treatment planning for HNC is considered a complex and labor-intensive process because this site typically includes a large number of organs-at-risk (OARs) for which the functional outcome after treatment is correlated with the radiation dose. These OARs are generally located in close proximity to the planning target volumes (PTVs) and should therefore be spared from radiation doses as much as possible. The main investigations and findings of the different chapters are summarized below.

During intensity modulated radiotherapy (IMRT), patients are typically treated using multiple static fields that individually deliver inhomogeneous dose distributions, which only when combined lead to sufficient and homogeneous irradiation of the tumor. These inhomogeneous dose distributions are created by moving the leaves of a multileaf collimator (MLC) through the treatment field, which only permits radiation to be delivered through the MLC aperture. During volumetric modulated arc therapy (VMAT) treatments, the treatment field rotates around the patient at the same time as the individual MLC leaves are moved in and out of the field to create different apertures and modulate the treatment beam. The positions and movement of the MLC leaves for IMRT and VMAT treatments are determined during an optimization process in which the planner sets optimization goals for the delineated structures. The optimization algorithms attempt to achieve MLC leaf

configurations that, as far as possible, result in a photon fluence / dose distribution that satisfies the optimization goals. Since the optimal treatment plan is often unknown before starting the planning process for individual patients, these optimization goals should be adjusted regularly during the treatment planning process to achieve high quality plans.

Previous investigations have shown that volumetric modulated arc therapy (VMAT) can lead to plans of similar quality as intensity modulated radiotherapy (IMRT), but with the advantage of offering reductions in delivery time and decreasing monitor unit (MU) requirements. VMAT is therefore increasingly being used for the treatment of HNC. **Chapter 2** showed that improvements in plan quality can be obtained by increasing the number of VMAT arcs. Additional arcs increase the number of possible MLC configurations, and these can be exploited to further decrease OAR doses and improve PTV dose homogeneity. This, however, comes at the expense of a longer beam-on time and an increased amount of MUs. It should also be noted that the results discussed in **Chapter 2** may be specific to Eclipse, as this treatment planning system (TPS) attempts to generate plans in which the gantry is rotating at maximum speed. Other TPSs allow for a slow-down of the gantry and, if this is fully exploited in the optimization algorithms, increasing the number of arcs may not necessarily improve plan quality.

Automated treatment planning solutions have the potential to both increase the efficiency in modern radiotherapy departments, and to improve resulting plan quality and consistency. Since different treatment planning techniques may lead to varying levels of OAR sparing and PTV dose homogeneity, the trade-offs between these parameters was investigated in **Chapter 3**. An exponential relationship was found, showing that small decreases in PTV dose coverage and homogeneity may lead to substantial improvements in OAR sparing, and vice versa. This should be taken into account when benchmarking manually and automatically created treatment plans that present varying levels of PTV dose homogeneity, since this may influence the achievable levels of OAR sparing. Despite this, however, there remain differences between major trial groups concerning desirable PTV dose criteria. In **Chapter 4**, plans were therefore created while satisfying the PTV dose coverage and homogeneity values specified by three different HNC protocols. It was found that using the planning guidelines as specified by the European Organization for Research and the Treatment of Cancer (EORTC) instead of those of the Radiation Therapy and Oncology Group (RTOG) could, on average, improve mean dose values to the salivary glands and swallowing muscles by 7.6Gy and 7.0Gy, respectively. While such mean dose reductions can have a large effect in the anticipated normal tissue complication probability (NTCP) parameters for the respective OARs, no comparison of outcome parameters have as of yet been reported between these trial groups to see if the stricter PTV dose criteria lead to better tumor control.

A technical account of the automatic interactive optimizer (AIO), which was developed as a part of this thesis, is provided in **Chapter 5,** along with a preliminary investigation of VMAT plan quality using ten HNC patients. AIO was designed to be used in conjunction with the interactive optimization algorithm of the Eclipse TPS. Before AIO was introduced, interactive optimization for HNC was a labor-intensive process, requiring the planner to continuously adapt the position of the dose-volume objectives relative to the dose-volume histogram (DVH)-lines that are displayed during the optimization process. AIO automates the adaptions of the objectives by scanning the optimization window, determining the position of the DVH-lines, and taking over control of the mouse cursor from the user. Automating the positioning of the optimization objectives leads to more consistent and frequent adaptions, ensuring that adequate attempts at sparing are made for all designated OARs throughout the optimization. A more in-depth clinical evaluation of AIO plan quality was provided in **Chapter 6** using seventy HNC patients and twenty locally advanced lung cancer (LC) patients. Since the clinically relevant OARs and their doses are different between both groups, the LC patients required a different approach to interactive optimization than HNC. This work demonstrated versatility of the AIO solution. AIO was introduced in clinical practice at the department of radiation oncology of the VU University Medical Center in February 2014 and has since then been used to automatically optimize all clinical HNC treatment plans. This has resulted directly from the research presented in this thesis. We have also shown that AIO typically results in more consistent sparing of the salivary glands and swallowing muscles, while its use improves planning efficiency by allowing the planners to invest their time in other high-value tasks while the treatment plan is being optimized.

Chapter 7 benchmarked VMAT plans for HNC created by RapidPlan against their manually optimized counterparts. RapidPlan is a commercial knowledge-based planning (KBP) solution developed by Varian Medical Systems (Palo Alto, USA). By using a model that correlates the geometric and dosimetric features of a library of previously created patient plans, RapidPlan can predict the achievable OAR DVHs for individual patients. These predictions can be used as input for the subsequently performed optimization process, thereby reducing the need for iterative adjustments to find optimal placement of the optimization objectives. For most OARs, RapidPlan improved mean dose values over the manually optimized plans, although model size and composition could influence these results. In addition, dosimetric outliers in the model library, e.g. plans that provide sub-optimal OAR sparing compared to the bulk of the model, may reduce the accuracy of the DVH predictions. **Chapter 8** therefore investigated the effect on plan quality when including

varying numbers of dosimetric outliers, in the form of plans in which no attempt was made to spare the salivary glands), to a 70-patient HNC model. These particular RapidPlan models were found to be robust against such outliers, up to 20 of which (out of a total of 70 plans) could be included in the library before affecting the quality of the resulting plans.

KBP solutions can also be used to provide quality assurance (QA) of plans, for example by comparing the achieved OAR doses with the doses that were predicted by the KBP model based on the patient's geometric features. This application was discussed in **Chapter 9,** where the accuracy of OAR dose predictions from a ninety patient RapidPlan HNC model were evaluated by comparing the predicted and achieved mean OAR doses for a twenty patient evaluation group. Excellent linear correlations were found between these parameters for all included OARs, although it was noted that the OAR dose predictions were less accurate in the high and low dose regions. As suggested in this chapter, KBP solutions could provide clinical trials a fast and patient specific method to evaluate whether submitted plans are deemed of sufficient quality.

The aforementioned investigations can be viewed as traditional treatment planning studies, in which different planning techniques are benchmarked using the same set of evaluation patients. This ensured that the results were not influenced by potential differences between the selected patients. Based on such studies, preferred techniques are often introduced in the clinic. However, the controlled environment in traditional planning studies makes it uncertain whether the predicted gains have been successfully translated into routine clinical practice. To investigate this, **Chapter 10** provided a longitudinal evaluation of routine clinical HNC plans delivered at the department of radiation oncology at the VU University Medical Center over the past decade. For four different periods, 30 patients were selected. Each period was characterized by a distinct approach to radiotherapy dose delivery and /or treatment planning. The transition from the first period (including manually optimized IMRT plans from 2005-2008 solely attempting parotid gland and spinal cord sparing), to the last period (automatically optimized VMAT plans from 2014-2015 that attempted sparing of the parotid glands, submandibular glands, oral cavity and a large number of swallowing muscles) was associated with substantial gains in OAR sparing. These gains neither degraded PTV dose homogeneity values, nor did they increase dose deposition in the remainder of the body volume. Since different patients were included in the different periods, statistical corrections were used to account for inevitable geometric differences between the groups.

Finally, proton therapy has often been suggested for the treatment of HNC because the steep dose fall-off after the proton beams' dose maximum (the Bragg-peak) offers the potential to improve OAR sparing over conventional photon techniques. The critical positioning of the Bragg-peak, however, may also make proton plans less robust for errors in patient positioning and range uncertainties (i.e., errors in Hounsfield density of the planning CT-scan). **Chapter 11** therefore investigated OAR sparing and plan robustness for different single- and multi-field optimized proton plans and compared these with VMAT photon plans. Plan robustness was evaluated by introducing ±3mm setup errors and ±3% range uncertainties to the plans and evaluating the resulting V95% and V107% values for the clinical target volumes. Multi-field optimized plans resulted in improved OAR sparing over single-field optimized proton plans and VMAT plans, but they were also less robust.

Different methods to achieve high quality IMRT and VMAT treatment plans for HNC were investigated in this thesis. Exploring optimal plan quality for individual patients has generally been limited to well-equipped and staffed, modern radiotherapy institutes. The advent of automated planning, however, may also facilitate the widespread creation of optimal plans in radiotherapy institutes that have fewer resources.

Samenvatting

De drie primaire behandelmethoden voor kanker zijn chirurgie, radiotherapie en chemotherapie. Voor een specifieke patiënt bepalen veel factoren, waaronder de locatie en stadium van de tumor, en de leeftijd, gezondheid en persoonlijke voorkeur van de patiënt, de keuze van behandeling, of combinatie van behandelingen. Radiotherapie speelt een belangrijke rol in de behandeling van hoofdhals kanker (HHK) en is voor de meerderheid van HHK patiënten minstens even effectief als chirurgische verwijdering van de tumor, met het functionele behoud van de organen in het hoofdhals gebied als belangrijk voordeel.

Voor radiotherapeutische behandelingen dient een bestralingsplan te worden gemaakt. Dit bestralingsplan moet ervoor zorgen dat de tumor met voldoende dosis wordt bestraald, terwijl het gezonde weefsel een zo laag mogelijke dosis krijgt. In de huidige klinische praktijk worden bestralingsplannen doorgaans gemaakt met behulp van computeralgoritmes waarmee het plan wordt geoptimaliseerd via een iteratief proces waarin met steeds kleiner wordende verbeteringen het uiteindelijke plan wordt bereikt. De kwaliteit van dit resulterende plan wordt echter sterk bepaald door de bekwaamheid van de planner en de keuzes van de behandelend arts, waardoor er grote verschillen kunnen ontstaan tussen de plannen van verschillende planners en verschillende radiotherapieafdelingen. Daarom is in dit proefschrift onderzocht hoe het maken van bestralingsplannen (*treatment planning*) voor radiotherapeutische behandelingen van HHK kan worden geautomatiseerd, zodat de invloed van individuen wordt weggenomen en tegelijkertijd getracht wordt de plannen verder te verbeteren.

Treatment planning voor HHK is een complex en tijdrovend proces omdat er in het hoofdhalsgebied een groot aantal organen (*organs-at-risk*, OARs) aanwezig zijn die schade kunnen ondervinden door een te hoge stralingsdosis. Deze OARs liggen doorgaans in de buurt van de doelgebieden (*planning target volumes*, PTVs) en zullen daarom zo goed mogelijk moeten worden gespaard. Hieronder volgt een samenvatting van de algemene onderzoeksvragen en belangrijkste uitkomsten die gepresenteerd zijn in de verschillende hoofdstukken van dit proefschrift.

Tijdens *intensity modulated radiotherapy* (IMRT) worden patiënten doorgaans bestraald met meerdere statische velden, welke individueel inhomogene dosisverdelingen leveren, en waarvan alleen de combinatie leidt tot voldoende en homogene bestraling van de tumor. Deze inhomogene dosisverdelingen worden gemaakt door tijdens de bestraling metalen bladen (*leaves*) van een *multi-leaf collimator* (MLC) door het veld te bewegen, zodat de straling alleen door de openingen van de MLC komt. Tijdens *volumetric modulated arc therapy*

(VMAT) behandelingen draait het bestralingsveld met een constante snelheid om de patiënt heen, terwijl de MLC *leaves*, net zoals in IMRT, tegelijkertijd worden bewogen. Geschikte posities en bewegingen van de MLC *leaves* voor IMRT en VMAT behandelingen worden bepaald door middel van een optimalisatieproces waarin de gebruiker optimalisatiedoelen aangeeft voor de verschillende ingetekende structuren. Door middel van een iteratief proces probeert het computeralgoritme MLC configuraties te vinden waarmee er zo goed mogelijk aan de optimalisatiedoelen wordt voldaan. Omdat bij het starten van het planningsproces het meest optimale bestralingsplan doorgaans niet bekend is voor individuele patiënten, moeten deze optimalisatiedoelen tijdens het *treatment planning* vaak worden aangepast om tot het best mogelijke plan te komen.

Eerdere onderzoeken hebben aangetoond dat met IMRT en VMAT kwalitatief gelijkwaardige bestralingsplannen kunnen worden gemaakt. VMAT wordt echter in toenemende mate verkozen boven IMRT voor de behandeling van HHK, onder meer vanwege de kortere bestralingstijd. Het onderzoek in **Hoofdstuk 2** heeft aangetoond dat VMAT bestralingsplannen verder kunnen worden verbeterd wanneer er meer dan twee bogen (*arcs*) voor de bestraling worden gebruikt. Hiermee wordt namelijk het aantal mogelijke oplossingen van de MLC verhoogd, waardoor de OARs beter kunnen worden afgeschermd van dosis en het doelgebied (PTV) homogener kan worden bestraald. Deze voordelen moeten echter worden afgewogen tegen de langere benodigde bestralingstijd. Mogelijk gelden de resultaten van **Hoofdstuk 2** alleen voor het treatment planningssysteem (TPS) Eclipse (Varian Medical Systems, Palo Alto, VS), omdat dit TPS bestralingsplannen tracht te maken waarin de bestralingsarm (*gantry*) op de hoogst mogelijke snelheid ronddraait. Met andere planningssystemen in combinatie met een bestralingsapparaat van een andere fabrikant kan de rotatiesnelheid van de *gantry* worden verlaagd, waardoor, mits dit volledig benut kan worden door de gebruikte optimalisatiealgoritmes, het toevoegen van extra *arcs* niet noodzakelijkerwijs zal leiden tot betere bestralingsplannen.

Het automatiseren van het *treatment planning* proces kan enerzijds zorgen voor een toename in de efficiëntie van moderne radiotherapieafdelingen, en anderzijds de resulterende plannen verbeteren en hun consistentie verhogen. Omdat de keuze van verschillende planningstechnieken kunnen leiden tot verschillen in de homogeniteit van de dosis in het PTV en de dosis naar de OARs, is de relatie tussen deze parameters onderzocht in **Hoofdstuk 3**. Het exponentiële verband dat werd gevonden laat zien dat wanneer de dosis in het PTV erg homogeen is, een kleine concessie hieraan kan zorgen voor grote verbeteringen in het sparen van de OARs, en vice versa. Dit is van belang wanneer handmatig en automatisch gemaakte bestralingsplannen met een andere PTV dosishomogeniteit worden vergeleken,

aangezien de laagst haalbare OAR dosis sterk wordt beïnvloed door de eisen die aan de PTV dosis gesteld worden. Desondanks zijn er grote verschillen tussen de PTV dosiscriteria van verschillende internationale onderzoeksorganisaties die multicenter studies doen met grote groepen patiënten. Daarom zijn er in **Hoofdstuk 4** plannen gemaakt die voldoen aan de PTV dekking en homogeniteitcriteria van drie verschillende planningsprotocollen voor HHK. Het bleek dat door het gebruiken van planningsrichtlijnen van de *European Organization for Research and the Treatment of Cancer* (EORTC) in plaats van planningsrichtlijnen van de *Radiation Therapy and Oncology Group* (RTOG), de gemiddelde stralingsdosis van de speekselklieren en slikspieren met respectievelijk 7.6Gy en 7.0Gy kunnen worden verminderd. Hoewel dit een groot effect kan hebben op de verwachtte toxiciteit, de *normal tissue complication probability* (NTCP) waarden voor de respectievelijke organen, zijn er tot op heden geen vergelijkingen gemaakt tussen de verschillende planningsrichtlijnen van deze organisaties om te onderzoeken of strengere PTV dosiscriteria leiden tot een betere controle van de tumor na de behandeling.

Als onderdeel van dit proefschrift is een methode ontwikkeld waarmee, zonder invloed van een planner, de interactieve optimalisatie in Eclipse automatisch gedaan kan worden. In **Hoofdstuk 5** wordt de technische ontwikkeling van deze *automatic interactive optimizer* (AIO) beschreven en de resulterende plankwaliteit onderzocht op tien HHK patiënten. Het handmatig interactief optimaliseren van HHK patiënten is een moeilijk en arbeidsintensief proces waarin de planner de optimalisatiedoelen van een groot aantal OARs moet blijven aanpassen afhankelijk van de dosis-volume histogram (DVH)-lijnen die worden getoond gedurende het optimalisatieproces. De AIO automatiseert deze handelingen door de DVH-lijnen in het optimalisatiescherm af te lezen, en de bewegingen van de cursor over te nemen van de gebruiker. De optimalisatiedoelen kunnen hierdoor vaker en consistenter worden aangepast, waardoor er ten alle tijden pogingen worden gedaan om de relevante organen voldoende van dosis te sparen. **Hoofdstuk 6** beschrijft een meer klinisch onderzoek, waarin AIO plannen voor zeventig HHK patiënten en twintig patiënten met lokaal gevorderde longkanker zijn vergeleken met plannen gemaakt door een ervaren planner. Omdat de klinisch relevante dosiswaarden voor de organen van deze patiëntengroepen verschillen, moest de AIO optimalisatie voor longkanker op een andere manier worden uitgevoerd dan voor HHK. De positieve resultaten van dit onderzoek hebben dan ook de veelzijdigheid van AIO gedemonstreerd. Sinds februari 2014 worden alle klinische bestralingsplannen voor HHK op de afdeling radiotherapie van het VU medisch centrum met behulp van AIO gemaakt, wat een direct gevolg is van het onderzoek dat is beschreven in dit proefschrift. AIO leidt, onafhankelijk van wie het plan maakt, tot meer consistente sparing van de

speekselklieren en slikspieren. Het gebruik van AIO is ook efficiënter aangezien de planners tijdens de optimalisatie hun tijd elders kunnen besteden.

In **hoofdstuk 7** is het gebruik van RapidPlan onderzocht voor de automatische optimalisatie van VMAT bestralingsplannen voor HHK. RapidPlan is een commercieel *knowledge-based planning* product ontwikkeld door Varian Medical Systems (Palo Alto, VS). Het gebruik van RapidPlan begint met het maken van een model waarin de geometrische en dosimetrische eigenschappen van een verzameling eerder gemaakte plannen wordt gecorreleerd. RapidPlan kan dit model gebruiken om, aan de hand van geometrische eigenschappen van een nieuwe patiënt, behaalbare DVHs van OARs te voorspellen. Deze voorspellingen kunnen worden gebruikt als startpunt voor het optimalisatieproces waardoor het iteratieve aanpassen van de optimalisatiedoelen om hun optimale plaatsing te vinden niet meer nodig is. De RapidPlan bestralingsplannen gaven voor de meeste OARs een lagere stralingsdosis dan in de klinische plannen, al speelde de samenstelling en grootte van het model een belangrijke rol in deze resultaten. Dosimetrische uitschieters (*outliers*) in de verzameling bestralingsplannen, bijvoorbeeld plannen waarin bepaalde organen slechter zijn gespaard vergeleken met de rest van de plannen in het model, kunnen de kwaliteit van het model en de nauwkeurigheid van de voorspelde OAR DVHs beïnvloeden. In **Hoofdstuk 8** zijn daarom verschillende aantallen dosimetrische *outliers* (bestralingsplannen waarin geen poging is gedaan om de speekselklieren te sparen) toegevoegd aan een RapidPlan model met zeventig HHK patiënten, waarna het effect op de resulterende plankwaliteit is onderzocht. De RapidPlan modellen bleken robuust tegen dit soort *outliers*, waarvan er tot 20 konden worden toegevoegd voordat de kwaliteit van de resulterende plannen werd verminderd.

Knowledge-based planning oplossingen, zoals RapidPlan, kunnen ook worden gebruikt om de kwaliteit van bestralingsplannen snel te beoordelen, bijvoorbeeld door de OAR dosiswaarden die behaald zijn tijdens *treatment planning* te vergelijking met de dosiswaarden die voorspeld worden op basis van de geometrie van de patiënt. Deze toepassing is besproken in **Hoofdstuk 9**, waar de nauwkeurigheid van de dosisvoorspellingen van een RapidPlan model met negentig HHK patiënten is onderzocht door de behaalde en voorspelde gemiddelde dosiswaarden te vergelijkingen voor een groot aantal OARs van twintig patiënten. Hoewel uitstekende lineaire verbanden werden gevonden voor alle OARs, konden de dosisvoorspellingen minder nauwkeurig zijn in de hoge- en lage dosisregio's. Zoals in **Hoofdstuk 9** is voorgesteld kunnen OAR dosisvoorspellingen in klinische trials dienen als snelle en patiënt specifieke methodes om de kwaliteit van ingediende plannen te beoordelen.

De voorgaande onderzoeken kunnen worden beschouwd als traditionele *treatment planning* studies, waarmee verschillende planningstechnieken worden vergeleken op dezelfde groep testpatiënten. De behaalde resultaten zullen hierdoor niet beïnvloed worden door mogelijke verschillen tussen de geselecteerde patiënten. Vaak worden nieuwe bestralingstechnieken aan de hand van dit soort traditionele planningsstudies in de kliniek geïntroduceerd voor de behandeling van toekomstige patiënten. Door de gecontroleerde manier waarop dit soort planningsstudies worden uitgevoerd is het echter niet zeker of de voorspelde winsten ook worden behaald wanneer de nieuwe planningstechnieken routinematig in de kliniek worden gebruikt. Om dit te beoordelen is er in **Hoofdstuk 10** een longitudinaal onderzoek uitgevoerd waarin klinisch gebruikte bestralingsplannen van de afdeling radiotherapie van het Vrije Universiteit Medisch Centrum uit het afgelopen decennia zijn vergeleken. Uit elk van vier verschillende periodes werden dertig bestralingsplannen geselecteerd. Elk van deze periodes werd gekarakteriseerd door een unieke bestralings- of planningstechniek. De eerste periode bestond bijvoorbeeld uit handmatig geoptimaliseerde IMRT plannen uit 2005-2008, waarin alleen geprobeerd werd de oorspeekselklier (glandula parotis) en ruggenmerg te sparen. De laatste periode bestond daarentegen uit automatisch geoptimaliseerde VMAT plannen uit 2014-2015, waarin tevens de onderkaakspeekselklieren (glandula submandibularis), mondholte en individuele slikspieren zo veel mogelijk werden gespaard. Bij het vergelijken van de bestralingsplannen bleek dat er in de laatste periode grote winsten waren behaald in het sparen van alle OARs vergeleken met de eerdere periodes. Bovendien kwamen deze verbeteringen niet ten koste van een minder homogene PTV dosis, noch van een toename aan dosis in de rest van het lichaam. Omdat de patiënten en hun geometrische verhoudingen anders waren in de verschillende periodes, waren statistische correcties toegepast om te compenseren voor deze geometrische verschillen.

Tot slot zou protonentherapie bestralingsplannen voor HHK kunnen verbeteren omdat de steile dosisafval na het dosismaximum van de protonenbundel (de zogeheten *Bragg-peak*) de mogelijkheid schept om OAR sparing te verbeteren vergeleken met de eerder beschreven fotonenbestralingstechnieken. De kritieke plaatsing van de *Bragg-peak* in de patiënt maakt de protonenplannen echter minder robuust tegen afwijkingen in patiëntpositionering en onzekerheden in de lokale dichtheid in een patiënt (met andere woorden, onzekerheden in de Hounsfield dichtheid van de CT-scan). In **Hoofdstuk 11** zijn daarom voor verschillende *single-field* en *multi-field* geoptimaliseerde protonenplannen de resulterende waarden van OAR sparing en robuustheid vergeleken met VMAT fotonenplannen. De robuustheid van de protonenplannen werd onderzocht door de dosisverdeling te simuleren bij ±3mm positioneringfouten en ±3% afstandsonzekerheden, en de resulterende dosiwaarden van

de doelgebieden te vergelijken. *Multi-field* geoptimaliseerde protonenplannen gaven verbeteringen in het sparen van de OARs ten opzichte van *single-field* geoptimaliseerde protonenplannen en VMAT, maar ze bleken ook minder robuust tegen positioneringfouten en afstandsonzekerheden.

In dit proefschrift zijn verschillende methodes onderzocht om de best mogelijke IMRT en VMAT bestralingsplannen voor HHK te maken. Tot op heden was het bereiken van zulke optimale plankwaliteit voor individuele patiënten voorbehouden aan moderne radiotherapie afdelingen. De komst van automatische *treatment planning* technieken kan er echter voor zorgen dat de best mogelijke plannen ook kunnen worden gemaakt in radiotherapie afdelingen die minder middelen tot hun beschikking hebben.

Dankwoord

Ik kijk met ontzettend veel plezier terug op mijn promotietraject. Dit komt voor een groot deel door de prettige samenwerking die ik heb gehad met mijn collega's. Ik wil de onderstaande personen dan ook hartelijk bedanken voor hun bijdragen.

Wilko, bedankt voor al je inzet en hulp die je mij de afgelopen jaren hebt gegeven. Je hebt mij al vanaf de eerste dag gestimuleerd om een kritische houding aan te nemen ten opzichte van de resultaten. Ik heb het ook zeer gewaardeerd hoe je innovatieve ideeën over het automatiseren van planning proces hebben gezorgd voor een brede toepassing van ons onderzoek. Bedankt voor je inzet om VAIO klinisch in gebruik te nemen. Dit gaf mij ook een belangrijke rol op de rest van de afdeling en demonstreerde perfect het belang van mijn onderzoek voor de kliniek.

Max, I have yet to meet someone that has a bigger drive for research and patient care than you. Thank you for all your help, guidance and support over the previous years. Thank you for the countless times you worked until deep in the night to revise our papers. Your passion for the work has been contagious, and I could not have wished for a better person to teach me about the clinical side of our investigations.

Patricia, ik heb erg genoten van onze samenwerking. Je inzet voor de KNO patiënten is noemenswaardig. Bedankt voor de vele avonden die je tot laat op de afdeling aanwezig was om slikspieren in te tekenen. Ik denk dat we trots mogen zijn op het eindresultaat; het is een heel mooi artikel geworden. Het ga je goed in Utrecht!

Prof. Slotman, Ben, bedankt voor alle steun die u mij de afgelopen jaren heeft gegeven. Het was zeer motiverend om op een afdeling te werken waar innovatie zo omarmd wordt, en ik hoop hier met mijn onderzoek aan te hebben bijgedragen.

Johan, bedankt voor alle steun en vertrouwen die je in mij had om dit promotieproject te beginnen, en bedankt dat ik nog even door kan gaan als post-doc!

Prof. Senan, Suresh, ondanks dat we niet hebben samengewerkt voor het beschreven onderzoek was u altijd goed op de hoogte van de ontwikkelingen. Bedankt voor alle feedback en hulp de afgelopen jaren.

Alex, the moment you came in for your Master internship work has gotten even more enjoyable. It was great to have a true friend at the workplace, and I'm proud of the awesome traditions we started, including IPA Friday, the occasional game of Frisbee, and more. Probably even more important, thanks for all your hard work in our investigations together.

Danique, bovenstaande geldt ook voor jou. Sinds je als masterstudent binnenkwam ben ik het werk alleen maar leuker gaan vinden. Ik was ontzettend blij dat je kon blijven als KLIFIO en ik heb genoten van onze samenwerking voor het protonen-paper. Bedankt voor alle gezellige kopjes 'roddel-'koffie en de nuttige werkgerelateerde discussies die we hebben gevoerd. **Tezontl**, bedankt voor de gezellige lunches, waarvoor ik de laatste paar maanden helaas minder tijd kreeg. De ESTRO feestjes zijn elk jaar weet iets om naar uit te kijken. **Mariët**, al hebben we voor de onderzoeken in dit proefschrift niet samen gewerkt, ik heb erg genoten van de samenwerking van afgelopen maanden voor het mamma scripting

Aan al mijn vroegere en huidige kantoorgenootjes, waaronder **Gwen**, **Naomi**, **Alex Louie**, **Sasha**, **Chin-Loon**, **Colien** en **Hilâl**, bedankt voor alle gezelligheid op de onderzoekskamer. **Colien** en **Hilâl**, heel veel succes komende tijd, ik heb er alle vertrouwen in dat jullie een geweldig resultaat gaan afleveren!

Bedankt aan alle **Planners** voor jullie geduld als VAIO weer eens vastliep. In het bijzonder wil ik graag **Omar Bohoudi, Ilonka** en **Roy** bedanken voor hun inzet voor VAIO en RapidPlan.

Omar Hertgers, bedankt voor je vriendschap en het meedenken met de cover, het ziet er geweldig uit!

Ilja, allereerst bedankt voor alle gezellige lunches (en belangrijker, de tik-ei competitie) toen je op de VU je stage liep. Onze vriendschap is alleen maar sterker geworden door de festivals op de Waddeneilanden, het tijdelijk zijn van Prinsengracht-homies, en de volleybalcompetities. Bedankt dat je als paranimf naast mij wilt staan op het podium.

Floris, ons 20-jarige vriendschap hebben we vorig jaar nog in Japan gevierd, hetzelfde land waar jij momenteel je droom (en carrière) aan het waarmaken bent. Bedankt voor alles deze afgelopen 20 jaar, maar in het bijzonder voor de game-sessies van afgelopen jaren, die waren perfect om even helemaal te ontspannen na een dag hard werken. Ik ben bijzonder vereerd dat je voor mijn promotie bent teruggekomen naar Nederland en naast mij wilt staan als paranimf.

Peter en **Jetske**, al vanaf mijn eerste dag samen met Laura hebben jullie mij met open armen verwelkomt. Ik zal jullie voor altijd dankbaar zijn hoe jullie mij een nieuw thuis in Amsterdam hebben gegeven. Ook voor alle gezellige etentjes en zaterdagochtendkopjeskoffie, bedankt!

Papa, zonder twijfel heb jij de basis gelegd voor mijn interesse in de wetenschap. Je innovatiedrift en 'er is altijd een oplossing'-instelling heeft mij altijd doen verder zoeken als de oplossing niet voor de hand lag. Ik stel het ook bijzonder op prijs dat je als leek mijn

artikelen bestudeerde, en ondanks alle vakjargon met zeer kritische vragen wist te komen. **Mama**, weinig zijn er zo zorgzaam als jij. Je immer positieve instelling heeft mij geleerd dat voor alle problemen wel oplossingen komen, en dat ik me nooit zorgen hoef te maken als het werk even niet zo lukte.

Ook de rest van mijn familie heeft een belangrijke rol gespeeld en heeft mij van jongs af aan onvoorwaardelijk gesteund. Grote broer **Bas** en zus **Susanne**, bedankt voor alles. Al was ik een nakomertje, jullie hebben mij altijd overladen met aandacht. Ook de rest van mijn familie, **Rob**, **Nel**, **Joop**, **Huan** en **Cliff** mogen niet ontbreken in dit dankwoord. **Oma**, u was altijd mijn grootste fan en sprak trots over uw kleinzoon en zijn promotieonderzoek. Het doet mij veel verdriet dat u het einde van mijn promotietraject niet mee heeft kunnen maken. Neefjes **Sylven** en **Sidney**, en nichtje **Zoë**, het is altijd een feest om terug naar Den Helder te komen en jullie verder te zien opgroeien.

Maar nog het meeste dankbaar ben ik jou, mijn liefste **Laura**. Woorden kunnen niet uitdrukken hoe ontzettend veel je mij hebt gesteund deze afgelopen jaren. Voor alles dat je mij in ons bijna achtjarig samenzijn hebt gegeven (en dat is heel erg veel), bedankt. Alles gaat ons voor de zon, en ik kan niet wachten om te zien wat de toekomst ons zal brengen.

Curriculum Vitae

Jim Tol was born in 1988 in Den Helder, the Netherlands. After finishing high school in 2007, he decided to enroll in the Bachelor of Theoretical Physics and Astronomy at the University of Amsterdam. During his studies, Jim became acquainted with the field of Medical Physics, particularly through an internship at the Department of Biomedical Engineering and Physics at the Academic Medical Center in Amsterdam. Inspired by the potential of Medical Physics to thoroughly improve the outlook of patients for a wide variety of diseases, Jim decided to continue his education with a Master's in Physics of Life and Health at the University of Amsterdam, which he finished Cum Laude in 2012. The required internship to finish the Master's involved the automatic segmentation of cardiac muscle fibers from cryomicrotome images, and gave him an important introduction to object-oriented programming.

Jim was introduced to the field of radiation oncology during his Master's through a literature study concerning novel dose delivery modalities in radiotherapy. The hand in hand relation between Physics and Medicine in this field fulfilled his professional interests and Jim therefore applied for a PhD research position regarding automated treatment planning strategies in the department of radiation oncology at the VU University Medical Center (VUmc), at which he started working in October 2012. Amongst other investigations that were performed during the PhD research, Jim created and evaluated his own software designed to automate the treatment planning process (VAIO – VUmc automatic interactive optimizer), and collaborated with Varian Medical Systems for the clinical evaluation of RapidPlan, a commercial knowledge-based planning solution. After finishing his PhD research, Jim has decided to continue as a postdoctoral researcher at the same department to continue the development and clinical introduction of automated treatment planning strategies.

List of publications

J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel. Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: an exponential trade-off with PTV dose homogeneity. Medical Physics **41**(2), 021722 (2014).

J.P. Tol, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel. Increasing the number of arcs improves head and neck volumetric modulated arc therapy plans. Acta Oncologica **54**(2), 283-297 (2015).

J.P. Tol, M. Dahele, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel. Different treatment planning protocols can lead to large differences in organ at risk sparing. Radiotherapy and Oncology **133**(2), 267-271 (2015).

J.P. Tol, A.R. Delaney, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel. Evaluation of a knowledge-based planning solution for head and neck cancer. International Journal of Radiation Oncology, Biology, Physics **91**(3), 612-620 (2015).

J.P. Tol, M. Dahele, J. Peltola, J. Nord, B.J. Slotman, W.F.A.R. Verbakel. Automatic interactive optimization for volumetric modulated arc therapy planning. Radiation Oncology **10**(1), 75 (2015).

D.L.J. Barten, **J.P. Tol**, Comparison of organ-at-risk sparing and plan robustness for spot-scanning proton therapy and volumetric modulated arc photon therapy in head-and-neck cancer. Medical Physics **42**(11), 6589-6598(2015).

J.P. Tol, M. Dahele, A.R. Delaney, B.J. Slotman, and W.F.A.R. Verbakel. Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? Radiation Oncology **10**(1), 234 (2015).

A.R. Delaney, **J.P. Tol**, M. Dahele, J. Cuijpers, B.J. Slotman, W.F.A.R. Verbakel. Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution. International Journal of Radiation Oncology, Biology, Physics **94**(3), 469-477 (2016).

J.P. Tol, M. Dahele, A.R. Delaney, P. Doornaert, B.J. Slotman, and W.F.A.R. Verbakel. Detailed evaluation of an automated approach to interactive optimization for volumetric modulated arc therapy plans. Medical Physics **43**(4), 1817-1829 (2016).

J.P. Tol, P. Doornaert, B.I. Witte, M. Dahele, B.J. Slotman, and W.F.A.R. Verbakel. A longitudinal evaluation of improvements in radiotherapy treatment plan quality for head and neck cancer patients. Radiotherapy and Oncology **119**(2), 337-343 (2016).